12-2014

Virtual Mobility Domains - A Mobility Architecture for the Future Internet

Hasan Tuncer

Virtual Mobility Domains - A Mobility Architecture for the Future Internet

by

Hasan Tuncer

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the B. Thomas Golisano College of Computing and Information Sciences
Rochester Institute of Technology

December, 2014

Signature of the Author _____

Accepted by _____
Director of the Ph.D. Program Date

B. THOMAS GOLISANO COLLEGE OF
COMPUTING AND INFORMATION SCIENCES
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Hasan Tuncer has been examined and approved by the dissertation committee as complete and satisfactory for the dissertation requirement for Ph.D. degree in Computing and Information Sciences

Dr. John F. Hamilton, Member Date

Dr. Sumita Mishra, Member Date

Dr. Andres Kwasinski, Member Date

Dr. Koushik Kar, Member Date

Dr. Nirmala Shenoy, Dissertation Supervisor Date

Dr. Hector Flores, Dissertation Examination Chairperson Date

To Türkmen Sultan

Curriculum Vitae

Hasan Tuncer was born in Denizli, Turkey on March 26, 1984. He attended the College of Engineering at Koc University, Istanbul, from 2002 - 2007 with full merit scholarship (ranked 112th among 1.5 million students in university entrance exam), where he obtained his Bachelor of Science degree in Computer Engineering (High Honors) in 2007. He began his doctoral studies in Computing and Information Sciences at the Rochester Institute of Technology in the Fall of 2007. He has been working at the wireless networking and security laboratory at Rochester Institute of Technology as a research assistant under the supervision of Professor Nirmala Shenoy since September 2007. His main research interests are in the area of mobility management for the future Internet architectures. Mr. Tuncer served as an adjunct professor teaching undergraduate level courses from 2009 - 2012 in the Department of Networking, Security, and Systems Administration at the Rochester Institute of Technology. Mr. Tuncer is a member of the Institute of Electrical and Electronics Engineers (IEEE) Communications Society. He served as session chair at the Scheduling Session of Ad-hoc and Sensor Networks Symposia at IEEE ICC 2012. He also served as peer reviewer for the 18 papers submitted to IEEE Transactions on Mobile Computing, IEEE Transactions on Wireless Communications, Elsevier Pervasive and Mobile Computing, and Wiley International Journal of Communication Systems; and to the following IEEE conferences: ICC, GLOBECOM, WCNC, ICCCN, ICNC, SoftCOM, and ISWCS. His publications resulting from his Ph.D. study include:

Journal Publications

1. H. Tuncer, S. Mishra, and N. Shenoy, A Survey of Identity and Handoff Management Approaches for the Future Internet, *Elsevier Computer Communications Journal*, Volume 36, Issue 1, 1 December 2012, Pages 63-79, ISSN 0140-3664.

2. H. Tuncer, A. Kwasinski, and N. Shenoy, Performance Analysis of Virtual Mobility Domain Scheme vs. IPv6 Mobility Protocols, *Elsevier Computer Networks Journal*, Volume 57, Issue 13, 9 September 2013, Pages 2578-2596, ISSN 1389-1286.

3. Y. Nozaki, H. Tuncer, and N. Shenoy, Evaluation of Tiered Routing Protocol in Floating Cloud Tiered Internet Architecture, submitted to *Elsevier Computer Communications Journal, Special Issue on Future Internet Testbeds* on June 2012. Currently in the first round of the review process.

Conference Proceedings

4. H. Tuncer, Y. Nozaki, and N. Shenoy, Virtual domains for seamless user mobility, in *Proceedings of the 9th ACM international symposium on Mobility management and wireless access, ser. MobiWac11*. Miami, FL, USA, 2011.

5. H. Tuncer, Y. Nozaki, and N. Shenoy, Virtual Mobility Domains - a mobility architecture for the future internet, *Communications (ICC), 2012 IEEE International Conference on*, vol. 2774, no. 2779, pp. 10-15, June 2012.

6. H. Tuncer, Y. Nozaki, and N. Shenoy, Seamless user mobility in Virtual Mobility Domains for the future internet, *Communications (ICC), 2012 IEEE International Conference on*, vol. 5860, no. 5865, pp. 10-15, June 2012.

7. H. Tuncer, N. Shenoy, A. Kwasinski, J. F. Hamilton, and S. Mishra, A novel user-centric handoff cost framework applied to the Virtual Mobility Domains and IPv6-based mobility protocols, *Global Telecommunications Conference (GLOBECOM 2012)*, vol. 2578, no. 2584, pp. 3-7, Dec. 2012.

8. Y. Nozaki, H. Tuncer, and N. Shenoy, A tiered addressing scheme based on a Floating Cloud internetworking model, in *Proceedings of the 12th international conference on Distributed computing and networking, ser. ICDCN11*. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 382393.

9. Y. Nozaki, H. Tuncer, and N. Shenoy, ISP tiered model based architecture for routing scalability, *Communications (ICC), 2012 IEEE International Conference on*, vol. 5817, no. 5821, pp. 10-15, June 2012.

Virtual Mobility Domains - A Mobility Architecture for the Future Internet

by

Hasan Tuncer

Submitted to
the B. Thomas Golisano College of Computing and Information Sciences
in partial fulfillment of the requirements
for the Doctor of Philosophy Degree
at the Rochester Institute of Technology

Abstract

The advances in hardware and wireless technologies have made mobile communication devices affordable by a vast user community. With the advent of rich multimedia and social networking content, an influx of myriads of applications, and Internet supported services, there is an increasing user demand for the Internet connectivity anywhere and anytime. Mobility management is thus a crucial requirement for the Internet today.

This work targets novel mobility management techniques, designed to work with the Floating Cloud Tiered (FCT) internetworking model, proposed for a future Internet. We derive the FCT internetworking model from the tiered structure existing among Internet Service Provider (ISP) networks, to define their business and peering relationships. In our novel mobility management scheme, we define Virtual Mobility Domains (VMDs) of various scopes, that can support both intra and inter-domain roaming using a single address for a mobile node. The scheme is network based and hence imposes no operational load on the mobile node. This scheme is the first of its kind, by leveraging the tiered structure and its hierarchical properties, the collaborative network-based mobility management mechanism, and the inheritance information in the tiered addresses to route packets.

The contributions of this PhD thesis can be summarized as follows:

- We contribute to the literature with a comprehensive analysis of the future Internet architectures and mobility protocols over the period of 2002-2012, in light of their identity and handoff management schemes. We present a qualitative evaluation of current and future schemes on a unified platform.

- We design and implement a novel user-centric future Internet mobility architecture called Virtual Mobility Domain. VMD proposes a seamless, network-based, unique collaborative mobility management within/across ASes and ISPs in the FCT Internetworking model. The analytical and simulation-based handoff performance analysis of the VMD architecture in comparison with the IPv6-based mobility protocols presents the considerable performance improvements achieved by the VMD architecture.

- We present a novel and user-centric handoff cost framework to analyze handoff performance of different mobility schemes. The framework helps to examine the impacts of registration costs, signaling overhead, and data loss for Internet connected mobile users employing a unified cost metric. We analyze the effect of each parameter in the handoff cost framework on the handoff cost components. We also compare the handoff performance of IPv6-based mobility protocols to the VMD.

- We present a handoff cost optimization problem and analysis of its characteristics. We consider a mobile user as the primary focus of our study. We then identify the suitable mwathematical methods that can be leveraged to solve the problem. We model the handoff cost problem in an optimization tool. We also conduct a mobility study - best of our knowledge, first of its kind - on providing a guide for finding the number of handoffs in a typical VMD for any given user's mobility model. Plugging the output of the mobility study, we then conduct a numerical analysis to find out optimum VMD for a given user mobility model and check if the theoretical inferences are in agreement with the output of the optimization tool.

List of Tables

List of Figures

List of Abbreviations

AAA	Authentication, authorization, and accounting
AR	Access router
AS	Autonomous system
BU	Binding update
BA	Binding acknowledgement
CAC	Common anchor cloud
CN	Correspondent node
CoA	Care-of address
DHCP	Dynamic host configuration protocol
FB	Forwarding base
FCT	Floating cloud tiered
HA	Home agent
HMIPv6	Hierarchical Mobile IPv6
HoA	Home address
IETF	Internet engineering task force
IP	Internet protocol
ISP	Internet service provider
LBU	Local binding update
LBA	Local binding acknowledgement
LCoA	On-link care-of address
LMA	Local mobility anchor
MAP	Mobility anchor point
MAG	Mobility access gateway

MIPv6 Mobile IPv6

MN Mobile node

OSI Open systems interconnection

PBU Proxy binding update

PBA Proxy binding acknowledgement

PMIPv6 Proxy Mobile IPv6

QoS Quality of service

RCoA Regional care-of address

RFC Request for comment

VMD Virtual mobility domain

Contents

*

Chapter 1

Introduction

The advent of the 1980s brought major changes in the Internet's operation, as commercial applications gained popularity. The number of devices and networks connecting to the Internet has increased along with the variety of applications and services. The unprecedented and significant technological advancements could not be foreseen during the initial design of the Internet. Despite these changes, the Internet continues to operate based on legacy principles and protocols. The inclusion of mobile devices and applications has affected the performance, scalability, and quality of service (QoS), which is due to factors such as mobile node handoff, re-addressing, routing, and security.

The current Internet requires efficient mobility management, which should provide a seamless mobility experience to users. Seamless mobility means low latency, low data-packet loss, and minimum quality of service degradation on an ongoing Internet session, while the mobile user moves from the coverage of one wireless access router to another, either in the same network or in different networks. The research effort of this dissertation aims at a novel mobility-management scheme that is capable of providing a seamless roaming experience to mobile users who are connected to the Internet, where the Internetworking model is one that has been designed for the future of the Internet.

The significant advances witnessed today are evidence of an era that was far ahead of the times when the current Internet was invented. Internet Engineering Task Force (IETF) initiatives developed mobility protocols such as Mobile IPv4 (MIPv4), Mobile IPv6 (MIPv6), Hierarchical Mobile IPv6 (HMIPv6), and Proxy Mobile IPv6 (PMIPv6) to support user mobility in the current Internet architecture as the demands for such services became significant. These protocols can be categorized broadly as network-based or host-based mobility protocols. They can also be categorized based on the mobility scope, such as macro-mobility protocols, dealing with the mobility of a user across administrative domains, and micro-mobility protocols handling the movement of a user across access

routers under the same administrative domain. Each type of mobility protocol provides a different way of defining mobility-management processes and structuring the mobile devices and network elements that support mobility. Though these protocols added mobility management to the Internet, problems in bringing an ideal, seamless mobility experience to users persisted because of the Internet architecture, which was not intended to support mobile users and the use of IP addresses, which are for the identification of a mobile node and routing packets to the mobile node.

In the current Internet architecture, the address of the mobile node changes when it moves to another access network. Therefore, the mobility protocols require address resolution in the new network, while the previous network is to be informed about the new address, by a process called *address binding*. Address binding starts with sending a binding update (BU) message that includes the new address of the mobile node to the home network. The home network then confirms the new address by sending a binding acknowledgement (BA) back to the mobile node [1]. The routing tables of the related nodes need to be updated to accommodate the packet routing to the new location of the mobile node, which is identified by its new address. These activities introduce latency and use additional computational resources (especially the wireless resources), which degrades performance. Other factors that need to be considered are the address length, because a long address uses up more of the costly wireless bandwidth, and the support for a large number of mobile nodes, because the number of wireless networks will increase in the future.

A seamless handoff experience requires less interruption in an ongoing session. Successful implementation of seamless mobility is closely related to the number of handoff-management messages between the wired and wireless devices, which handle the session, the number of the nodes that have to change their routing table entries (such as the mobile node, correspondent node, and routers in the previous and new networks), and the amount of the change in the current session setup to forward packets to the mobile node at its new access router. Using IP addresses for identification and routing results in high-mobility messaging and routing table updates, which increase the handoff latency and the signaling overhead. This could result in an interruption of the session and hence a service-quality degradation.

In this dissertation, we present the design and implementation of a novel, mobility architecture, called Virtual Mobility Domain (VMD), which works with Floating Cloud Tiered (FCT) internetworking model, which is proposed for a future Internet. We derive the FCT model from *the tiered structure* existing among ISP networks [2]. The resulting topological connections exhibit a hybrid structure that uses to its advantage the attributes of hierarchical and distributed structures. In this structure, there can be several entities

in one tier who operate in a distributed and autonomous manner. However, entities at a lower tier are customers of entities at a higher tier, exhibiting a hierarchical relationship. In the FCT Internetworking model, granularity and modularity were introduced to enable movement of entities across tiersindependent of their relationship with other entities [3].

Our novel mobility architecture defines various scopes of VMDs that can span several Autonomous Systems (ASes) or Internet Service Providers ISPs) to support both intra- and inter-domain roaming. The VMD architecture supports network-based mobility management by assigning a single address to a mobile node and, hence, limits the involvement of a mobile node in mobility management. The tiered structure in the FCT model is leveraged by the VMD to provide collaborative handoff management in a mobility domain that can span several networks. The VMD architecture is unique because of the collaborative network-based mobility management scheme that operates with the new tiered Internet working model; the structuring and the coordination of the network entities; and the usage of the inheritance information at the tiered addresses in packet routing. We conduct analytical studies and simulations to evaluate the performance of the VMD architecture in comparison with the current mobility protocols such as MIPv6, HMIPv6, and PMIPv6. We then did an analytical optimization study to find the optimum VMD that minimizes a mobile user's handoff cost depending on his mobility preferences and system parameters.

The remainder of this thesis is organized as follows. Chapter 2 includes the literature review of the mobile node identity and handoff-management covering mobile IP protocols and next-generation mobility solutions. We also provide a survey of mobility models and handoff-cost optimization studies to have all the related works for the upcoming chapters. The fundamentals of the proposed VMD architecture are presented in detail in Chapter 3 followed by the performance analysis of the VMD architecture in comparison to IPv6-based mobility protocols in Chapter 4. The handoff-cost framework implications are explained in Chapter 5. Optimization of handoff-cost considering mobile user as the primary focus is presented in Chapter 6. Finally, conclusions of this thesis and future work are discussed in Chapter 7.

Chapter 2

Literature Review

This chapter presents the review of the previous literature that is related to the research presented in the next chapters. First, a survey of the identity- and handoff-management solutions proposed in future Internet architectures are presented. Mobility protocols developed by the Internet Engineering Task Force initiatives are discussed to give the background on the user mobility support challenges with the current architecture. The next-generation network architectures supported by global initiatives are presented and analyzed in terms of their support for seamless user and device mobility. Furthermore, the survey is extended to include the architectures proposed for wireless mesh networks, which are envisioned to be a part of the next-generation networks with their self-organizing and self-configuring network characteristics.

The United States National Science Foundation's Future Internet Design Initiative [4] and Future Internet Architecture Project [5], the European Union's 6th and 7th Framework [6] Programs, the Asia Consortium [7], and New Generation Networks [8] in Japan supports evolutionary solutions to overcome the challenges encountered by the current Internet architecture. The Global Environment for Network Innovations [9] in the United States and the Future Internet Research and Experimentation [10] in Europe provide large-scale experimental network infrastructure for validation of new protocols and schemes.

Mobility is one of the challenges of existing and future networked applications and services. Future Internet design requires an understanding of the current status of mobility solutions, the approaches adopted by them, the challenges that they have targeted, and their limitations, given that they have to operate within the current Internet architecture.

[*] Portions of this chapter previously appeared as:
H. Tuncer, S. Mishra, and N. Shenoy, A Survey of Identity and Handoff Management Approaches for the Future Internet, *Elsevier Computer Communications Journal*, Volume 36, Issue 1, 1 December 2012, Pages 63-79, ISSN 0140-3664.

In this chapter, we provide the identity and handoff- management solutions for the current Internet architecture for a reader to have background information and then focus on a survey of future solutions from several global projects. Identity and handoff management are closely related topics because handoff management solutions get affected by how a mobile node is identified. Furthermore, a mobile node's address is used to trace it while it is moving and also to route the data packets to the mobile node on its new network. Other factors, such as application level solutions, QoS, and security provisioning, are also important in supporting seamless handoff management. However, to maintain our focus, they are not covered in this study.

One of the main goals of the new mobility architectures is to provide service that satisfies mobile users' needs such as handoff. Users' movements cause handoffs; hence, the study of mobile user movement is an important element while designing the new mobility architectures. User movement may depend on the roaming environment, such as urban, suburban, streets, highways, etc. Further, users may show independent, group-based, or random behaviors. The various roaming characteristics of mobile users are studied extensively in the literature. In Section 2.5, we provided an overview of the most common mobility patterns, such as random-walk mobility, fluid-flow mobility, nomadic-community mobility, and Manhattan-grid mobility models.

Current mobility protocols aim to provide seamless handoffs of mobile nodes. A seamless handoff requires low handoff latency and reduced data packet loss in an ongoing session, while a mobile node transfers from one access router to another in the same network or in another network. Successful implementation of seamless mobility is closely related to the number of handoff-management messages between wired and wireless devices that handle the session, the number of nodes that have to change their routing table entries (such as mobile node, correspondent node, and routers in the previous and new networks), and the change in the current session setup to forward packets to the mobile node at its new access router. There are various categories of protocols: network-based, host-based, micro-mobility, and macro-mobility. Each type of protocol provides a different way of organizing the network and handling the handoff management.

Mobile user's handoff causes signaling overhead, latency, location-tracking cost, and packet delivery cost. One of the mobility study goals is to decrease the costs associated with a handoff. In Section 2.6, we provide an overview of the literature on handoff-cost-optimization studies. Optimization studies focus on different mobility parameters. Most of them focus on the costs that affect service providers, such as packet delivery cost and location-tracking cost, while a few of them focus on the costs that affect mobile users, such as handoff delay and mobile device power consumption. The aim of these optimization studies is to adjust the network topology, or improve handoff-related processes depend-

ing on the network conditions and mobile user preferences.

This chapter surveys the literature over the period of 2002-2012 for mobile node identity and handoff management. In addition, this chapter provides literature review on mobility models and handoff-cost optimization. This chapter is organized as follows. In Section 2.1, the surveys published in the mobility research area are presented. Section 2.2 covers the mobility protocols developed under the Internet Engineering Task Force initiatives, followed by the solutions proposed toward future Internet architectures in Section 2.3. The discussion of the approaches in Mobile IP protocols and next-generation mobility solutions is provided in Section 2.4. Mobility models are explained in Section 2.5. Last, the handoff-cost-related optimization studies are presented in Section 2.6. Concluding remarks are presented in Section 2.7.

2.1 Mobility Architecture Related Works

In the last 15 years, user mobility support has been researched extensively. There have been tens of evolutionary or revolutionary mobility approaches proposed to provide better handoff management in IP networks, cellular networks, ad-hoc networks, or wireless mesh networks. Likewise, there have been several literature reviews covering these mobility protocols and focusing on different aspects of mobility management. In this section, we aim to present all the relevant surveys that investigate these protocols.

Akyildiz et al. [11] give a qualitative comparison of the mobility protocols running in all-IP-based wireless systems, categorizing them as network layer, link layer, and cross-layer approaches. In [12], a detailed comparative analysis of location update, handoff latency, and signaling overhead performance of the mobility architectures and protocols are presented. Furthermore, handoff management, paging, scalability, and robustness of some mobility protocols are examined in [13]. Sun and Sauvola [14] present the limitations of Mobile IP in solving the micro mobility challenges, and the possible solutions to address these challenges are examined in [15, 16].

Xie and Wang [17] investigate the handoff management in wireless mesh networks while IP mobility protocols' deployment in mobile ad-hoc networks (MANETs) is examined in [18]. In [19, 20], the authors provide a survey of integration of 3G and wireless local area network (WLAN), focusing on underlying network architectures, handoff management, and QoS.

In [21], a taxonomy and survey of location management strategies applied by mobility protocols are given. El Maliki et al. [22] cover identity management approaches, or standards considering privacy and security aspects. Furthermore, the requirements for

Internet mobility and a review of the primary handoff support methods used by IP or cellular network protocols are presented in [23].

Conti et al. [24] present the research challenges towards the design of the future Internet such as scalability, robustness, security, energy efficiency, and flexibility. Future Internet design initiatives focusing on network virtualization, network management, routing, resource sharing, optical networking, and security are explained in [25]. The architectural, socio-economic, and security approaches in European next generation network projects are discussed in [26].

Our work is distinct from existing surveys described in this section because (i) it provides a comprehensive analysis of future Internet architectures and mobility protocols in light of their identity and handoff management schemes; and (ii) it presents a qualitative evaluation of current and future schemes on a unified platform.

2.2 IP-based Internet Architectures

Since Internet's advent and the first request for comment (RFC) published by the Internet Engineering Task Force in April 1969, it has come a long way, where today several millions of networks and billions of devices connect to the Internet. The number of wireless devices connecting to the Internet however has far exceeded the number of wired devices. The inevitable need for the Internet connectivity on the move eventually required the development and deployment of networking protocols to support handoff of a mobile node. A mobile node's mobility is categorized based on mobility scope: macro mobility and micro mobility. While macro-mobility refers to the mobility across administrative domains, micro-mobility refers to the movement of the user across access routers or base stations under the same administrative domain. Furthermore, depending on the access technologies that a mobile node is handing off between, categories of vertical and horizontal handoff exist. In vertical handoff, a mobile node moves between different network types such as IEEE 802.11 WLAN to 3G cellular network while in horizontal handoff, a mobile node moves between same type of access networks such as WLAN to WLAN, or 3G network to 3G network etc.

Handoff process can be broken down into the following steps regardless of the mobility scope and the access network technology [27]:

1. *Handoff Initiation:* The decision of handoff request to a new network is made considering several criteria. In horizontal handoff, few of the criteria are received signal strength, signal to noise ratio, bit error rate, and channel availability. In vertical handoff, additionally, battery lifetime, available bandwidth, latency and congestion in the network, network coverage, mobility characteristics of a mobile node,

number of users in the network, policies and billing constraints are also taken into account before handoff. Combinations of these metrics are standardized as media independent handoff function module in IEEE 802.21 [28]. IEEE 802.21 provides a shim layer between Open Systems Interconnection (OSI) layer 2 and layer 3 for helping in handoff initiation, decision, and execution by coordinating the exchange of information between 802.3, 802.11, 802.15, 802.16 and 3G networks.

2. *Handoff Decision:* Handoff decision process is categorized as network-controlled handoff, mobile-assisted handoff, and mobile-controlled handoff. In network-controlled handoff, the network handles the necessary measurements and handoff decision, while in mobile-assisted handoff , a mobile node makes the measurements and waits for a network's decision on handoff. However, in mobile-controlled handoff, a mobile node decides when to handoff based on the measurements made by both the mobile node and the network.

3. *Handoff Execution:* Handoff is executed as either hard or soft handoff. In hard handoff, also called break-before-make, the ongoing connection with a current network is broken first, then a connection with a new network is made. In soft handoff, also called make-before-break, a mobile node is connected to both networks at the same time and hands off to the new network completely after all the mobility related processes are completed.

IETF aims to standardize different network types such as WLANs under 802.11 a/b/g/n, mesh networks under 802.11s, wireless personal area networks under 802.15, IPv6 over Low power WPAN that works with 802.15.4, broadband wireless access - WiMAX under 802.16. As stated before, IETF also introduced 802.21 to provide a framework for media independent handover focusing on OSI layer 2. Seamless mobility management maintaining the desired QoS provisions is a challenging task because of the change in network connection, access technology, network condition, and mobile node identifiers. Current research aims to provide solutions focusing on different OSI layers. For instance, IETF introduces 802.21 focusing on OSI layer 2, MIPv4, MIPv6, Fast MIPv6, HMIPv6, and PMIPv6 focusing on OSI layer 3, and -further, the Session Initiation Protocol (SIP) as signaling protocol for controlling voice and video sessions focusing on OSI application layer.

In this survey, we concentrate on OSI layer-3 identity and handoff management because a mobile node needs to preserve its identity regardless of its network point of attachment, and supporting handoff between different networks is vital for providing seamless mobility experience and maintaining the QoS provisioning for the user. Furthermore, in the current Internet architecture, its IP address is used to trace the mobile node and also to route the data packets to it in its new network. In this section, the identity and handoff management design fundamentals of MIPv4, MIPv6, Fast MIPv6, HMIPv6, and PMIPv6 mobility protocols are presented. Knowing the phases that Mobile IP went

through helps in understanding the reasons behind the current challenges of the Internet and the design philosophy of the future Internet projects discussed in Section 2.3.

2.2.1 Mobile IP for Internet Architecture

Mobile IP was first designed as an extension to the IPv4 protocol, and it was named as MIPv4 [29]. Subsequently MIPv6 [1] was proposed when IPv6 was introduced to overcome the limitations of IPv4.

Identity Management in Mobile IPv4/IPv6

Mobile IP uses IP addresses to identify a mobile node. Mobile IP also uses IP addresses to locate the mobile node, and to forward packets destined to a mobile node via its IP address. However, a mobile node acquires a new IP address called care-of address (CoA) from a foreign network, where it is roaming. Before using the address, a mobile node has to do duplicate address detection to check the uniqueness of the new address in MIPv6 [30].

To handle this address change issue, Mobile IP uses the home address of a mobile node (HoA) in its home network as its global identifier. A mobile node is thus expected to register its care-of address to its home network. Home network deploys home agent (HA) which handles registration of mobile node's care-of address at the home network. After a mobile node gets the care-of address, the mobile node and home agent exchange binding update and binding acknowledgement messages. The home agent is also responsible of forwarding the packets to the mobile node using the mobile node's care-of address.

Handoff Management in Mobile IPv4/IPv6

As a mobile node continues to use its home IP address as a global identifier, a correspondent node does not have to be aware of a mobile node's current care-of address. The home agent intercepts these packets on behalf of the mobile node and then forwards data packets to the mobile node using IP-in-IP packet encapsulation or tunneling [31]. However, packets that are sent from the mobile node are not handled in this way, but are instead sent straight to their destination. Hence, this packet routing process is called *triangular routing* in MIPv4 [32]. This non-optimal packet routing and tunneling however impose a high redirection load on the home agent and cause handoff latency as well. Therefore, MIPv6 introduced route optimization to overcome this issue.

MIPv6 with *route optimization* enables a mobile node to communicate directly with a correspondent node using its care-of address without a home agent intervention. The Route optimization process requires a sequence of signaling message exchanges between

a mobile node, a home agent, and a correspondent node as depicted in Fig. 2.1 [33].

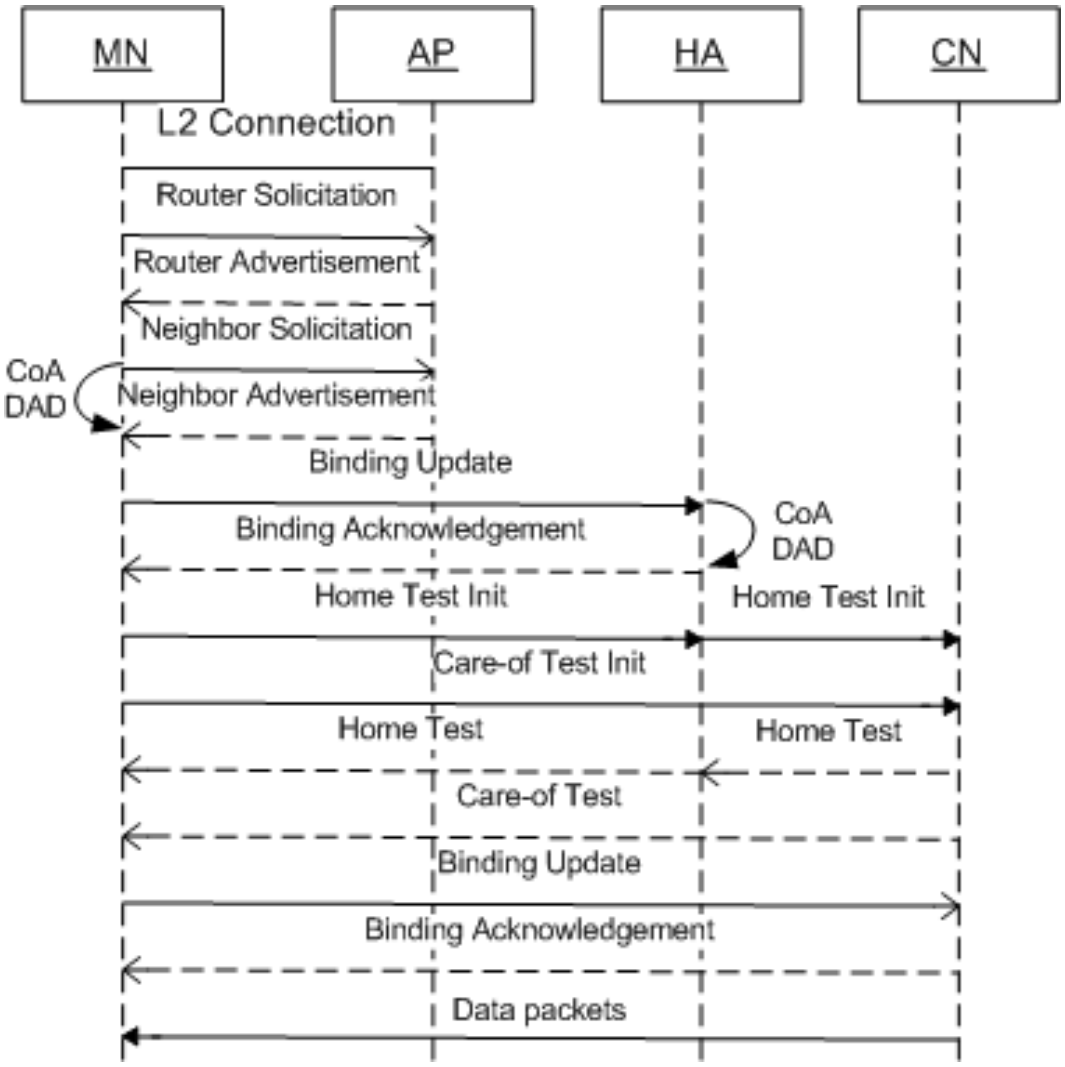


Figure 2.1: The message flow for MIPv6 with route optimization when mobile node moves to a new network.

As it can be observed in MIPv4 and MIPv6, mobile node starts communicating with a home agent after it gets a care-of address from a foreign network which introduces some latency. To overcome this problem, Internet Engineering Task Force proposed Fast MIPv6 as an extension to MIPv6. Fast MIPv6 allows a mobile node to establish a new temporary care-of address before breaking its connection with its old access router which

is called *anticipated handoff* [34]. When the mobile node is attached to the new access router, it can continue its communication with its new already-known address. If the anticipated handoff fails, the mobile node can always carry out a traditional handoff process. Moreover, Fast MIPv6 sets up a tunnel between the old access router and the new access router for the transmission of the data packets buffered at the old access router during the handoff process.

2.2.2 Hierarchical Mobile IPv6 (HMIPv6)

HMIPv6 [35] was designed to provide seamless handoff management for a mobile node within an administrative domain. Therefore, it is categorized as a micro-mobility management protocol. When the mobile node is roaming within an HMIPv6 domain, it does not have to send binding update messages to the home network or the correspondent node. HMIPv6 reduces the signaling load in the network by managing handoff locally in the domain and is thus more scalable and can support more mobile nodes. Fig. 2.2 illustrates a typical HMIPv6 deployed network.

Identity Management in Hierarchical Mobile IPv6

HMIPv6 uses two addresses to support the mobile node's micro-mobility. On-link care-of address (LCoA) is created based on access router link and a regional care-of address (RCoA) is created based on currently connected network's prefixes [36]. On-link care-of address is local identifier for a mobile node within a domain while regional care-of address is used to identify a mobile node globally.

Handoff Management in Hierarchical Mobile IPv6

HMIPv6 introduces a concept of mobility anchor point (MAP) that manages a micro-mobility of a mobile node within a domain. The mobile node exchanges local binding update (LBU) and local binding acknowledgement (LBA) messages with the mobility anchor point to register to a new access router. The mobile node then sends a binding update message to its home agent and correspondent node for them to bind the regional care-of address with the home address of the mobile node. If the mobile node moves within the same mobility anchor point domain as in Fig. 2.2, its regional care-of address will not change. The mobile node has to only register its new on-link care-of address to the mobility anchor point. This is one of the advantages of using micro-mobility protocols over macro-mobility protocols, because in a macro-mobility protocol, whenever a mobile node changes its address, a home agent has to be updated, which results in higher signaling load, increased latency, and eventually more packet loss.

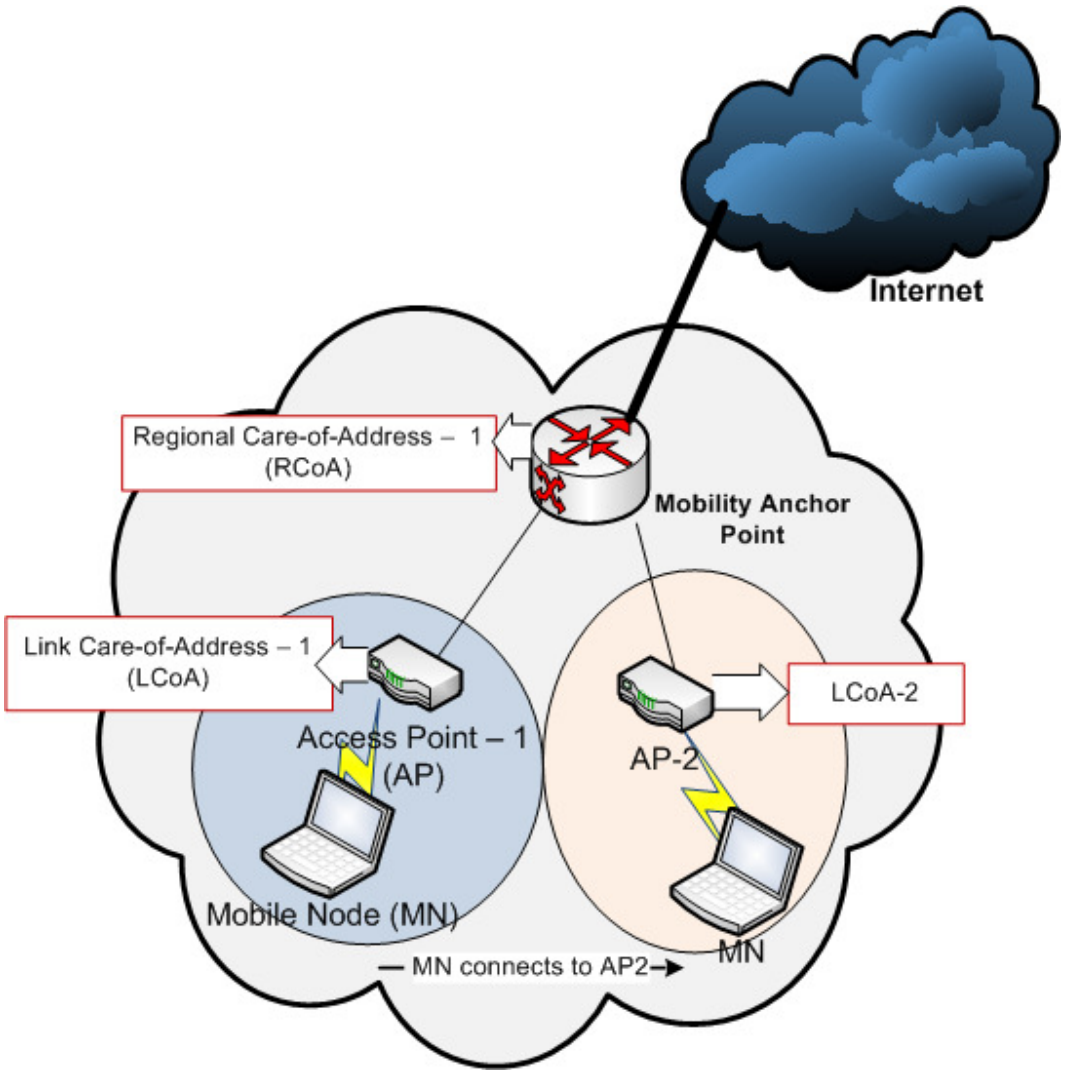


Figure 2.2: Mobility anchor point deployment in a typical HMIPv6 network and the addresses mobile node acquires in the HMIPv6 domain.

When the correspondent node or the home agent have packets to send to the mobile node, they will address the packets to the mobile node's regional care-of address. Then, the mobility anchor point intercepts these packets and sends them to the mobile node through a bidirectional tunnel binded to the on-link-care-of address of the mobile node.

2.2.3 Proxy Mobile IPv6 (PMIPv6)

PMIPv6 was proposed by the Network-based Localized Mobility Management Internet Engineering Task Force Working Group [37]. PMIPv6 is a network-based micro-mobility management protocol, and it does not require a mobile node to incur any mobility related signaling such as sending of binding updates, and encapsulation/decapsulation of data packets [38]. This is unlike MIPv6, HMIPv6 and Fast MIPv6 which propose host-based solutions and require a mobile node to actively involve in handoff management processes.

Identity Management in Proxy Mobile IPv6

PMIPv6 identifies a mobile node with a 128-bits long IPv6 address in a new network. Mobile node is not involved in address creation process and this address does not change as long as mobile node moves within the same PMIPv6 domain [39].

Handoff Management in Proxy Mobile IPv6

PMIPv6 introduces a mobility access gateway (MAG) module which is installed on access routers and a local mobility anchor (LMA) which is a wired node that all access routers have connection to. Mobility access gateway has the main role of detecting the mobile node's movements and initiating mobility-related signaling via local mobility anchor on behalf of a mobile node [40]. Local mobility anchor module decides on the mobile node's new address and then enables mobile node to communicate with external network nodes. To do that, mobility access gateway establishes a bidirectional tunnel with local mobility anchor.

As illustrated in Fig. 2.3, once a mobile node attaches to a mobility access gateway module for the first time, the mobility access gateway and the local mobility anchor exchange proxy binding update (PBU) and proxy binding acknowledgement (PBA) messages, and confirm the mobile node's profile with an AAA server. The local mobility anchor sends an address assigned to a mobile node via a proxy binding acknowledgement message. The local mobility anchor also sets up a bidirectional tunnel with the mobility access gateway for the mobile node to be able to communicate with a correspondent node.

To compare the current and future mobility protocols under a unified platform, a set of categories is identified (see Appendix A for details). Section 2.4 gives a qualitative discussion of all these protocols under this developed platform.

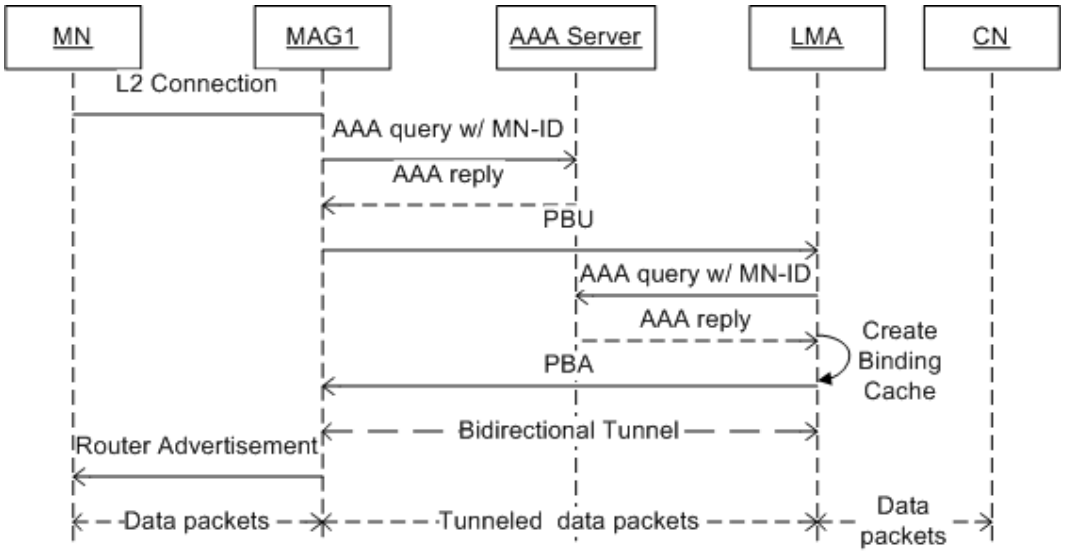


Figure 2.3: The micro-mobility related message flow occurs between mobility access gateway and local mobility anchor in PMIPv6.

2.3 Next Generation Mobility Solutions

Mobility protocols built to operate on the current IP architecture are discussed in the previous section. The performance of these protocols in terms of total overhead, handoff latency, capability of handling high mobility traffic and their limitations are studied in [39, 41–43]. The global initiatives stated at the beginning of the chapter are also indicative of the realization of the need for new Internet design approaches. In this section, we present the mobile node identification and handoff management approaches from the projects supported by the aforementioned initiatives. In Section 2.4, Tables 2.3, 2.4 and 2.5 present the features of these protocols with respect to selected categories such as mobility scope, handoff management, target network, mobile node address etc. The following projects and protocols are covered: MobilityFirst, eXpressive Internet Architecture (XIA), Ambient Networks, Designing Advanced network Interfaces for the Delivery and Administration of Location independent, Optimized personal Services (DAIDALOS), AKARI, Host Identity Protocol (HIP), Internet Indirection Infrastructure (i3), Host Identity Indirection Infrastructure (Hi3), Locator Identifier Separation Protocol (LISP), Mobility and Multihoming Supporting Identifier Locator Split Architecture (MILSA), CARrier grade MEsh Networks (CARMEN), HURRICANE, and MobileNAT.

2.3.1 MobilityFirst

The MobilityFirst Project [44], funded by National Science Foundation Future Internet Architecture program, involves eight universities from the United States. The aim of the project is to address mobility, multihoming, connectivity robustness, context-aware routing, and security challenges found in the current Internet architecture. To overcome these challenges, MobilityFirst follows key design principles such as separation of names from addresses, decentralized naming service, and generalized delay tolerant network (GDTN) with storage-aware routing. In order to maintain the focus of this survey, we will not go into details of every component and functionality; instead we will explore how host/node identification and handoff management are handled by MobilityFirst.

Identity Management in MobilityFirst

MobilityFirst provides three levels of identification as depicted in Fig. 2.4 [45]. At the highest level, each entity, e.g. computing devices, sensors and multimedia content, is presented with human-readable, context strings such as "Joe's laptop" or "Movie-A". At the second level, these entities are specified using a long-term, globally unique ID (GUID) from a flat naming space which does not depend on a network attachment point. Finally at the third level, these entities are specified with a network address such as IP address. The proposed architecture provides a decentralized naming service which has the following components: (i) Name Certification Service maps a human readable name to a GUID; (ii) The packet headers will have both a GUID and network addresses that will be protected by public key cryptography; and (iii) Location Service maps a GUID to a complete network address that is used for routing [44]. This hierarchical addressing solution allows the routing design to address mobility and varying level of connectivity considering mobile nodes and their associated applications as first-class Internet citizens.

Handoff Management in MobilityFirst

Handoff is handled using two mechanisms depending on the mobility characteristics of user. First, if a mobile node moves slowly, the location service is updated to reflect the new network address. Second, to provide ongoing session continuity, a home agent is deployed to redirect traffic to the new address of the node. MobilityFirst architecture deploys a delay-tolerant routing with in-network storage [46, 47]. In case of rapid host mobility, the network can use late or repeated binding to resolve a GUID to a network address at different points along the route [45].

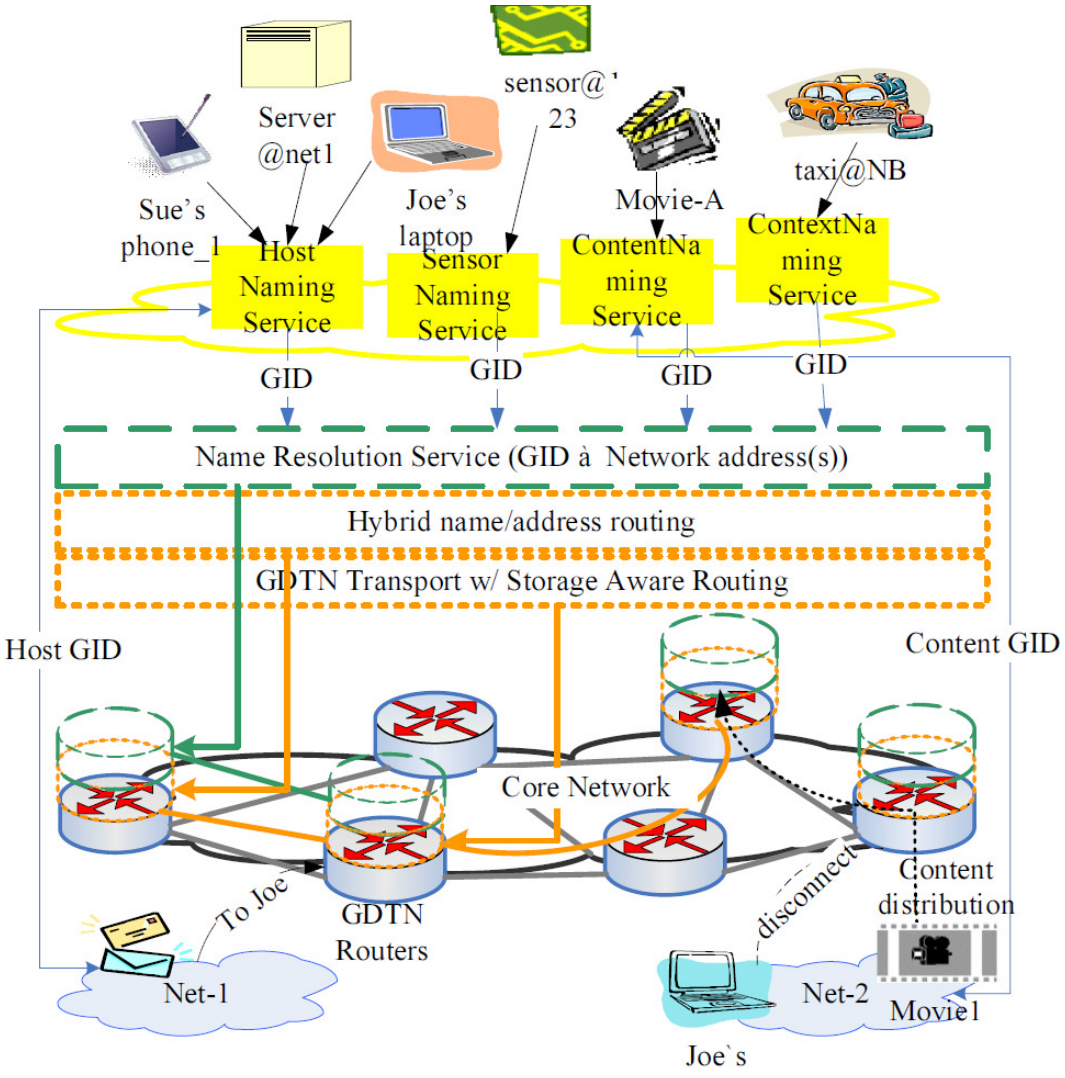


Figure 2.4: MobilityFirst protocol stack [45].

2.3.2 eXpressive Internet Architecture (XIA)

eXpressive Internet Architecture (XIA) [48] is also funded by National Science Foundation Future Internet Architecture program. XIA aims to preserve the strengths of current Internet architecture while substantially improving security, and building in the ability to support evolving network functionality over time. XIA introduces a new protocol called XIP as a replacement for IP which introduces new protocol stack, rich addressing and per-hop forwarding semantics [49].

Identity Management in eXpressive Internet Architecture

XIA defines hosts, services, contents, and administrative domains with unique eXpressive identifiers (XIDs). The eXpressive identifiers are mapped to locators using either a naming service e.g., Domain Name System or based on a locally maintained mapping to internal address e.g., Medium Access Control (MAC) address for a bluetooth device.

Handoff Management in eXpressive Internet Architecture

The building block of XIA for mobile users is called Tapa [50]. XIA provides segment based routing where each segment can be very diverse, ranging from wireless access networks such as multi-hop mesh networks, 802.11, and bluetooth to wired segments in the Internet or an enterprise network [50]. Each segment is responsible for delivering data from one end of the segment to the other end. If a mobile node changes its administrative domain, the mobile node's eXpressive identifier does not change in the new administrative domain. If there is any ongoing session, it is transmitted by XIP routers in the old administrative domain to the mobile node. For the new communications, a new mobile node address is created by prepending the eXpressive identifier of the new administration domain to the eXpressive identifier of the mobile node.

2.3.3 Ambient Networks

Ambient Networks [51], a large-scale collaborative project supported by the European Union 6th Framework Program, was set up for investigating future communication mechanisms. This project aimed to create a complete and coherent wireless network solution based on dynamic composition of networks through an instant establishment of inter-network agreements. The concept offers common control functions to a wide range of different applications and access technologies, enabling the integrated, scalable, and transparent control of network capabilities. Even though the project is currently closed, all the developed concepts are available in the repository.

Ambient Networks consists of three distinct components. The connectivity component abstracts existing network infrastructure on top of which the Ambient Networks

functionality resides. The interface component is for managing resources such as routers, switches, relays; providing application and service independence; and enabling interaction between different network technologies. Finally, the control space component provides naming framework, connectivity abstractions, security architecture, multi radio access, resource management, QoS, congestion control, mobility control functions, smart multimedia routing, and transport protocol for service-specific overlay networks, context awareness function, dynamic business agreement establishment and execution functions, and plug-and-play support functions.

Identity Management in Ambient Networks

The proposed architecture adopts a layered naming model to provide dynamic indirection between names, addresses, and identities which are currently being used [51]. Ambient Networks offers two connectivity abstractions: bearer and flow. Bearer provides abstraction to application services, or a specific data objects i.e. Session Initiation Protocol services and web pages through Ambient service interface which includes application point of attachment that can be compared with TCP/IP socket API. Bearer also provides the end-to-end customized transport service that supports all the functionality provided by the control space. On the other hand, flow is abstraction of the connectivity provided by the underlying technology where network nodes, mobile nodes, links, and paths reside. Flow is constrained to a single network technology and addressing domain. The naming in Ambient Networks is illustrated in Fig. 2.5 adopted from [51].

Handoff Management in Ambient Networks

In Ambient Networks, handoff management is handled by several subcomponents. When a mobile node enters a new network, *mobility triggering management* subcomponent collects and identifies triggers that include quality of service information, user policy information, security information, and end-to-end path information from different sources [52]. A handoff decision is based on the information retrieved from these triggers in conjunction with the *multi radio resource management* subcomponent. Next, a handoff mechanism to be used for the mobility event is selected from the *handoff toolbox* [53]. Handoff mechanism is selected primarily based on the types of endpoints that the mobility event affects and the performance requirements for flow continuity. Finally, at the handoff execution stage, the mobile node changes its network point of attachment. Hence, the mapping between its network and application points of attachment is updated at the mobile node. The actual updated mapping is defined by the mobility protocol that is being used to support the handoff. The network layer update is performed by a mobility protocol such as MIP or HIP, selected from the toolbox. This requires interactions with the *multi-radio resource management* subcomponent to identify and modify affected flows. This handoff execution process is controlled by *handoff and locator management* subcomponent which

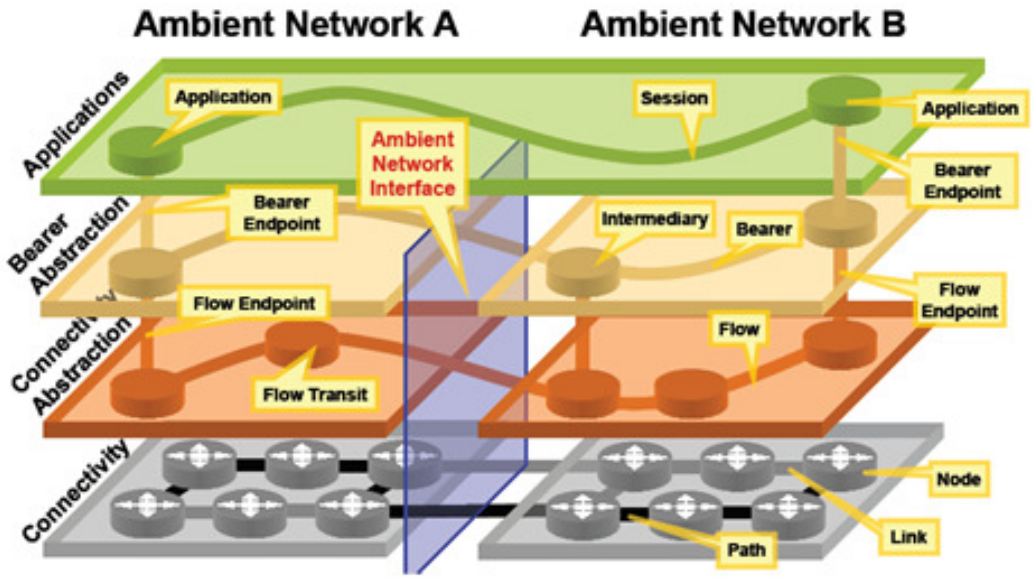


Figure 2.5: Connectivity abstractions in Ambient Networks [51].

actually aggregates procedures to include handoffs between access routers within a single radio network, between different access technologies, between different IP address spaces, multiple service provider domains, or application level handoffs. Furthermore, *reachability management* subcomponent enables a correspondent node to initiate communication with a mobility endpoint regardless of its current location.

Ambient Networks' mobility control space and handoff toolbox components enables the integration of many different standards such as GSM, UMTS, Mobile IP, Host Identity Protocol, SIP [54], Stream Control Transmission Protocol [55], distributed hash table based handoff management and so on. This feature aims to bring maximum mobility support between networks based on existing and future mobility protocols.

2.3.4 Designing Advanced Network Interfaces for the Delivery and Administration of Location independent, Optimized Personal Services (DAIDALOS)

Designing Advanced network Interfaces for the Delivery and Administration of Location independent, Optimized personal Services, DAIDALOS [56], is supported by European Union 6th Framework Program and has 46 collaborators from industry and academia. The DAIDALOS vision is to provide secure, personalized services built on seamlessly integrated heterogeneous network technologies including cellular, satellite, broadcast,

wired/wireless networks, and sensor networks.

Identity Management in DAIDALOS

DAIDALOS architecture supplies Virtual Identity (VID) Framework in which a profile of an entity (single user or group of users) may stem from contracts with different networks and services. Subsets of this entity profile are called entity profile views, that are the virtual IDs of the entity. A user can choose the virtual identity - service provider mapping. After virtual identity is confirmed by the service provider, the entity gets an IP address tied to that virtual identity [57]. Virtual identity concept requires ID-Broker, that supplies entity's location to correspondent node and proxies the request to the entity and ID-Manager. ID-Manager provides interface for creating, managing, and destroying virtual identities by abstracting entity's physical interfaces.

DAIDALOS also provides Virtual MAC infrastructure, which enables an entity to have two or more virtual identities bind to one physical interface to be able to access different providers. These virtual identities can be expanded to the relationships between banks, governmental institutions, operators, and service providers.

Handoff Management in DAIDALOS

DAIDALOS defines mobility as users can change their device while remaining connected to the Internet; the device can change its point of connection; the session can be moved from one interface to another in the same device; or the source of the service can change during the session. DAIDALOS splits the architecture into local and global domain to support mobility of a mobile node. In global domain, the mobile node's macro-mobility is managed by MIPv6, or HIP while the mobile node's micro-mobility in local domain is managed with HMIPv6 and PMIPv6 with support of different access technologies such as WLAN, WiMAX, 3GPP LTE and AD HOC/NEMO [58]. When the mobile node connects to a new network, depending on the mobility protocol, it gets a care-of address associated with the virtual identity that mobile node had. Depending on the low or high privacy concerns, home address and care-of address can be independent or dependent relatively. If high privacy is required, an entity receives a service via home address or care-of address which are associated with one virtual identity rather than several virtual identities. The dependency of mobile node's care-of address and home address to virtual identity causes an increased latency due to the fact that a new virtual identity needs to be bootstrapped every time when a mobile node moves to a new domain and creates a new care-of address. Furthermore, care-of addresses are by definition a locator of the point of attachment of the mobile node and hence creating a correlation between virtual identity and care-of address may be challenging.

2.3.5 Host Identity Protocol (HIP)

Host Identity Protocol [59–61] created by Bob Moskowitz in 1999, is in the process of standardization by HIP Internet Engineering Task Force Working Group [62]. HIP offers a method of separating the end-point identifier and locator roles of IP address to ease mobility and multihoming as well as security.

Identity Management in HIP

HIP separates the endpoint identifier and locator roles of IP address by introducing a one 128 bits long host ID tag (HIT). Host ID tag is public key of a public-private key pair. HIP also creates thin layer between the IP layer and the transport protocols. Applications are bound to host ID tag while IP still acts as a locator. The binding between IP and host ID tag happens at the kernel. The two communicating nodes at HIP are confident about each other's host ID tag after a four-way handshake, called base exchange. Base exchange employs Diffie-Hellman authenticated key exchange method, illustrated in Fig. 2.6. Then, these two HIP-aware end point communicate with each other using their host ID tags in a secure way by having encapsulated security payload Security Associations.

Handoff Management in HIP

HIP enables mobile node mobility across IPv4 and IPv6 [63, 64]. HIP provides handoff management by splitting mobile node identifier and locator. Mobile node executes base exchange mechanisms with correspondent node. Then two HIP-aware-end nodes, correspondent node and mobile node, start communicating using their host ID tags. Network layer connectivity goes over IP addresses but upper layers uses host ID tags. If the mobile node moves to a new network, its IP address changes. Even transport layer connectivity does not get affected because it relies on host ID tag, the correspondent node has to be informed about the mobile node's IP address change. Therefore, the mobile node executes end-to-end three-way UPDATE signaling mechanism [65] in which the mobile node sends UPDATE message to the correspondent node with its new address and security association generated during the base exchange. Then, the correspondent node sends an UPDATE acknowledgement. The mobile node evaluates the UPDATE acknowledgement and then echoes the nonce in the UPDATE acknowledgement message back to the correspondent node. After the process is completed, the mobile node continues its communication with the correspondent node.

HIP rendezvous servers (RVSs) [66] and HIP local rendezvous servers [67] propose extensions to HIP by deploying rendezvous servers for enhanced macro and micro-mobility management respectively. Rendezvous server is an initial contact point for mobile node and provides the HIP services to mobile node. Mobile node registers its new IP address and host ID tag to rendezvous server. Furthermore, mobile node uses IP address

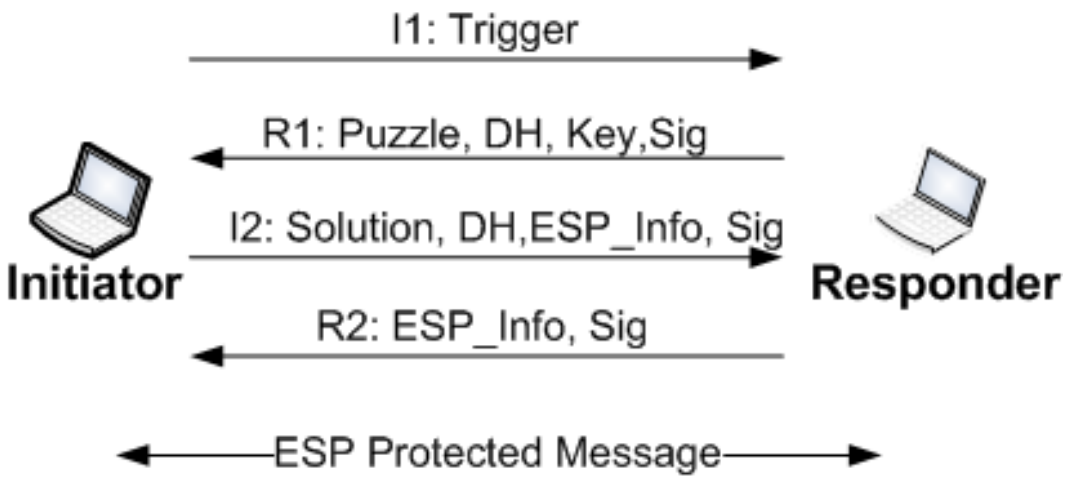


Figure 2.6: HIP base exchange mechanism.

of rendezvous server in base exchange with correspondent node. Then, correspondent node sends packets to the IP address of rendezvous server with the host ID tag of mobile node. The rendezvous server forwards the packet to mobile node's actual IP address. If mobile node changes its location, hence IP address, the transport layer connectivity with correspondent node does not break down because host ID tag stays same. At the network layer, correspondent node continues to sends packets to IP address of rendezvous server. Mobile node only needs to update its IP address at the rendezvous server for rendezvous server to be able to direct packets to mobile node's new location. [68] proposes improvements on the localized micro-mobility management by enhancing the functionality of Local rendezvous server. HIP requires that all the changes happen in the end-hosts which may potentially require significant changes to the current Internet structure and could lead to compatibility issues for existing protocols and applications.

2.3.6 AKARI

AKARI [69], supported by Japanese government, aims for a future Internet architecture to serve demands of solving societal challenges and the conditions of future available technologies. Some of the proposed approaches include ID/locator split, cross layer design, control layers with different time-scale behaviors, optical access switching and optical paths, overlay network, network virtualization, support for seamless movement across variety of wireless access technologies, and packet division multiple access for wireless connectivity assuring quality of service.

Identity Management in AKARI

AKARI applies ID/locator split approach to give mobility and multihoming support to a larger number of users and devices across dynamic heterogeneous environments. Node identifiers can be totally independent of network topology and internetworking technology. Some identifiers might have global scope while others might be private and they might be tied to more than one locator. Application and transport layers use string or bit stream to identify communicating nodes. Identifiers have two versions: names and IDs. AKARI deploys identity management servers (IMS) to assign locally unique names to the nodes. Each identity management server has a string identifier such as "mynetwork.com" and a node receiving an address from identity management server can have a name e.g. "my.pc". Global names are created combining local name and identity management server name. On the other hand, ID is a hash value of a name. IDs are included into packet header to identify the source and destination nodes. Locators might be global or local and one locator might have more than one IDs. Identity management server stores dynamic information such as mapping between names, IDs, and locators. Furthermore, AKARI also deploys a name mapping server (NMS) to store mapping between identity management server and locators which do not change so often. The ID layer is inserted between network and transport layer. If a mobile node wants to communicate with another node, first it gets the ID of that node from identity management server and then gets the locator from location management server. While the ID-locator mapping system is kept at the edge network to provide fast mobility support, global locator information is kept at the core network to have scalable routing. For transition purposes the first 64 bits of IPv6 address is proposed to be used as an ID and the remaining bits can be used as a locator.

Handoff Management in AKARI

In AKARI, each local access network is connected to the Internet via gateways. If a mobile node moves to another network, MIPv6 is deployed to inform a correspondent node about the mobile node's locator address change. However, the transport layer connectivity will not be affected from the mobile node's mobility since it is tied to the mobile node's ID. One of the challenges for AKARI is the support of micro-mobility.

2.3.7 Internet Indirection Infrastructure (i3)

i3 [70] offers rendezvous-based overlay indirection service to provide robust, scalable, and efficient system for handoff management, multicast, and anycast communications on the Internet [71].

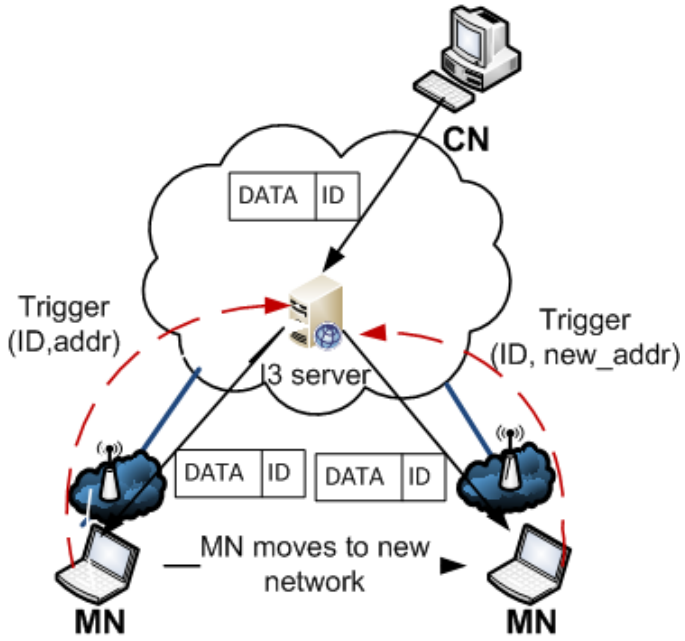


Figure 2.7: i3 handoff management mechanism and data communication scheme.

Identity Management in i3

i3 uses IDs which are simple integer numbers and addresses which are IPv4/v6 addresses depending on the network of a mobile node.

Handoff Management in i3

Let's first analyze how i3 provides communication between the nodes. i3 deploys i3 servers that keep the node ID and address. If a mobile node wants to be reachable by a correspondent node, it sends a trigger which includes its ID and actual IP address (ID, addr) to i3 servers. i3 servers store these triggers. If the correspondent node wants to send a packet to the mobile node, it sends the packet with the destination ID, (ID, data). The i3 server receives the packets and then forwards the packets to the address of the matched ID owner. If the mobile node changes its network, it only has to send a new trigger, which contains the same ID but the new address (ID, new_addr) to i3 servers as illustrated in Fig. 2.7. Then i3 servers forward packets to the new address of the mobile node. If sender caches the address of i3 servers, it achieves fast packet delivery. Zhuang et al. [72] build robust overlay architecture for mobility (ROAM) on top of i3 which controls the placement of i3 servers in i3 to provide efficient routing, fast handoff, personal/session mobility, and location privacy. End-hosts can use off-line heuristics to choose triggers that

are stored at i3 servers close to either itself or the correspondent node to avoid triangular routing problem as well as location privacy. Location privacy can also be achieved if the mobile node advertises two triggers: (ID, ID') to the i3 server near to correspondent node and (ID', addr) to the i3 server near itself. During the movement, the mobile node can do soft handoff via multicasting triggers with new address before moving to the new network. i3 heavily relies on i3 servers and hence the location of the servers should be considered carefully to provide fast, reliable, and scalable mobility service.

2.3.8 Host Identity Indirection Infrastructure (Hi3)

Hi3 [73–75] integrates HIP and i3 to provide better seamless mobility support and security. Hi3 inherits mobility, multihoming, and basic security mechanisms from HIP. Hi3 also deploys i3's secure integrated overlay rendezvous infrastructure as a control plane.

Identity Management in Hi3

Hi3 uses IP addresses as locator for a mobile node. On the other hand, the mobile node can have two identifiers: host ID tag is used as a public (server identifier) and ID is used as private (lower naming layer) identifier. Host ID tag is used to create association between a client and a server and then the communication between server and client continues on private identifiers. It results in performance increase and security improvement because correspondent node talks to the mobile node directly with the private trigger not the public one.

Handoff Management in Hi3

When a mobile node moves to a new network, it only sends triggers to a naming server to update its address, as in i3. This process introduces less signaling overhead compared to HIP. Hi3 outperforms i3 and enhances flexibility and security compared to HIP.

2.3.9 The Locator Identifier Separation Protocol (LISP)

The Locator Identifier Separation Protocol (LISP) [76,77] adopts a locator/ID split approach and a network-based map-and-encapsulate scheme [78] to solve naming/addressing, mobility, and multihoming challenges. LISP runs on IPv4 and IPv6 architectures as an incremental protocol which can be used for IPv6 transition, improving traffic engineering, and reducing size of core routing tables [79].

Identity Management in LISP

In LISP, a mobile node has endpoint identifier (EID) and routing locator (RLOC). LISP deploys mobility anchor point servers [80] to store endpoint identifier - routing loca-

tors mapping. It also introduces two network nodes: ingress tunnel router (ITR) and egress tunnel router (ETR). Ingress tunnel router performs endpoint identifier to routing locator look up and encapsulates the packet with the routing locators for both source and destination address fields separately. Egress tunnel router decapsulates the accepted packet. Furthermore, LISP Alternative Topology (LISP-ALT) [81] builds an overlay logical topology running instance of Border Gateway Protocol (BGP) [82] typically over GRE tunnels to ease endpoint identifier to routing locator mapping process. LISP also inserts Map-Encap layer into network layer of OSI to ease the mapping of endpoint identifier with routing locator and encapsulation of the packets with routing locator. LISP handles mapping at LISP -ALT control plane and it handles encapsulation and tunneling at data plane.

For example, in Fig. 2.8, a mobile node with endpoint identifier 2.0.0.2 wants to send a packet to a destination which has endpoint identifier 3.0.0.3. The packet is retrieved by ITR1. If ITR1 does not know endpoint identifier to routing locator mapping for 3.0.0.3, it encapsulates the packet with the outer header having source address (routing locator of ITR1) and destination address, endpoint identifier 3.0.0.3. The data probe is sent into the LISP-ALT topology. The packet follows the paths computed by BGP in the LISP-ALT topology to ETR1. ETR1 decapsulates the packet and forwards the inner packet to 3.0.0.3. Then, ETR1 also sends a mobility anchor point reply (MAP-reply) to ITR1 which tells that endpoint identifier-routing locator mapping for 3.0.0.3 has ETR1 whose routing locator is 12.0.0.2. After ITR1 receives the MAP-reply, it encapsulates the packets with its own address as a source and ETR1's address as destination address, and sends over the Internet.

Handoff Management in LISP

LISP provides routing scalability, mobility, and multihoming support with the help of naming mobile node with endpoint identifier and routing locator as well as deploying Map-Encap layer. When a mobile node changes its connection, the mapping between the endpoint identifier and the routing locator has to be changed. This process will cause delay and packet drop [83]. For the other mobility related issues, LISP gets benefit of the IPv4/6 mobility protocols.

2.3.10 Multihoming Supporting Identifier Locator Split Architecture (MILSA)

Mobility and Multihoming Supporting Identifier Locator Split Architecture, MILSA [84, 85], adopts a hybrid design of ID/locator split and core-edge separation concepts. The aim is to provide a solution to the current Internet's challenges such as renumbering, routing scalability, mobility, and multihoming [86].

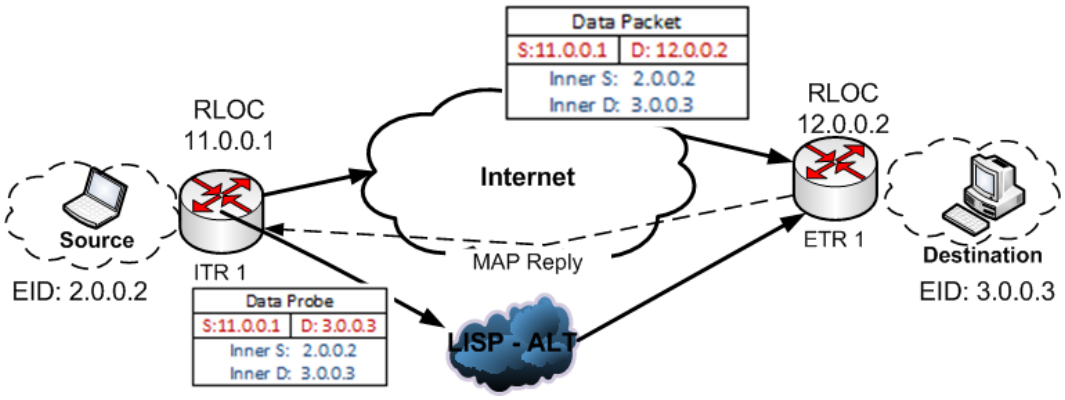


Figure 2.8: Naming and packet forwarding in LISP.

Identity Management in MILSA

In the MILSA communication environment, there are computers, mobile computing devices, firewalls, servers, humans, companies, departments, cities, and countries. MILSA defines realms as a hierarchical group of these objects that logically belong to the same organization such as an administrative domain [86]. Each of them has a name/ID, called hierarchical URI-like Identifier (HUI), and locator. Each name/IDs is valid within a realm. URI-like Identifier can be "{Hashed Key}.JohnRoberts.mail.us.google.com" which is 128-bits long. The first part is the hash of public key that uniquely identifies an object while the second part is hierarchical logical part. Locators are the IP addresses which are given by topologically aggregated physical network called zones. Hence locators are used for routing. MILSA deploys realm-zone bridging (RZB) servers that perform the mapping with the identifiers and locators. While an IP address can be used as a locator, they also propose a hierarchical code based locator structure as an alternate locator address scheme in Fig. 2.9. However, they did not go into details of how legacy routers in the zone will operate according to this addressing.

Service Provider Code	Country Code	Province Code	Region Code	End-host Code

Figure 2.9: Hierarchical code based locator structure used by MILSA.

Fig. 2.3.10 shows the format of the packets travel on the networks.

The length of the addresses in this packet structure may not be wireless friendly because wireless networks require less bandwidth usage.

Figure 2.10: Data packet format used at MILSA.

MAC	Next Hop Locator	Dst Locator	Src Locator	Dst HUI	Src HUI	Payload

Handoff Management in MILSA

If a mobile node changes a network, the mobile node's locator address changes while its name/ID does not change. MILSA inserts URI-like Identifier mapping sublayer between application layer and network layer which is similar to HIP as a concept. The mobile node informs realm-zone bridging servers about the locator change. MILSA deploys an access zone router and a backbone zone router that use Mobile IP protocols for network layer handoff management. The access zone router in the edge performs the identifier/locator indirection, and gets the ID/locator mapping from the realm-zone bridging servers. The access zone router routes the packets to the remote host through the backbone zone router [87].

2.3.11 Carrier Grade Mesh Networks (CARMEN)

CARMEN [88,89], CARrier grade MEsh Networks, aims to integrate heterogeneous wireless network technologies in a multi-hop fashion to provide scalable and efficient ubiquitous connectivity. CARMEN basically adopts a general IEEE 802.21 architecture but differs by targeting heterogeneous mesh networks in a media independent manner [90]. IEEE 802.21 focuses only on handoffs between heterogeneous technologies. The CARMEN cross-layer approach is composed of a MAC abstraction sub-layer and a mesh functions sub-layer. The MAC abstraction sub-layer is located between the subnet layer and the routing layer, in order to hide technology specific issues at the low layers. The Mesh functions sub-layer consists of routing, capacity handling, handoff management, self-configuration, and monitoring modules.

CARMEN requires a mobile node to be 802.21 compatible to use the services of the CARMEN mesh. CARMEN mesh point (CMP) is a node that is equipped with CARMEN capabilities. CARMEN access point (CAP) is a CARMEN mesh point with the capability to provide the mobile node's access to the CARMEN mesh. Typically, CARMEN access points have one or more access radio interfaces dedicated for the mobile node's use, and thus these interfaces do not carry traffic to other CARMEN mesh points. CARMEN gateway (CGW) is a CARMEN mesh point that also provides connectivity to the network provider's core or backbone network. CARMEN gateway is located at the boundary between the core network and the CARMEN mesh.

Identity Management in CARMEN

CARMEN requires unique 48-bits identifiers per node and per interface. CARMEN provides media independent handoff function to support mobility. In CARMEN, a mobile node gets two addresses. The first address is network access identifier encoded IP address i.e. IPv4 or IPv6 address or domain name for layer-3 communication. The second address is network access identifier encoded link layer address for layer-2 communications. Each interface ID is created converting the 48-bits interface ID into an extended unique identifier, 64-bits long. This interface ID is then used to generate the link local address via the IPv6 stateless address auto configuration mechanism defined in [36].

Handoff Management in CARMEN

The hierarchical positioning of CARMEN wired nodes provides easy management of macro-mobility (handoff between CARMEN gateways) and micro-mobility (handoff between CARMEN access points) of mobile nodes. Handoff management scheme provides the following functionalities available at CARMEN access point, CARMEN gateway, and core network: location registration/update, flow control, handoff preparation, deciding on target CARMEN access point, and execution. Before a mobile node is moving to a new CARMEN access point, all possible new CARMEN access points are investigated and then the resources of the chosen CARMEN access point are reserved. Then, the mobile node connects to the new CARMEN access point. If the new CARMEN access point is connected to a different CARMEN gateway, then the CARMEN gateway, in addition to CARMEN access point, is involved in handling the transfer of the ongoing session to the new network and handling the correspondent node information. CARMEN mainly presents the characteristics of network-based handoff management because of the high involvement of CARMEN access point and CARMEN gateway in mobility management.

2.3.12 HURRICANE

HURRICANE [91] supports ubiquitous and optimal broadband connectivity among cooperative networking environments. It aims to provide an optimized handoff operation across various radio cooperative networking environments such as GPRS/UMTS, Wi-Fi, WiMAX, and Digital Video Broadcasting-Handheld.

Identity Management in HURRICANE

HURRICANE does not implement protocol stacks from scratch; instead it specifies and modifies mechanisms that allow the existing protocols to operate in an efficient way. Therefore, the architecture does not propose a new naming/addressing scheme. It uses IPv4/IPv6 addressing scheme.

Handoff Management in HURRICANE

HURRICANE provides vertical handoff using the concept of IEEE 802.21 [92] and enhances it inspiring with IEEE 1900.4 [93]. HURRICANE deploys vertical handoff controllers, context information collectors (CIC), and handoff managers (HM) both at mobile node and network side. Furthermore, the media independent handoff service of HURRICANE hides the technological differences among radio access networks. These components collaboratively give mobility decision according to a user's quality of service requirements, policies, network resources, and network condition. Then, these components orchestrate the legacy handoff management protocols such as MIPv4, MIPv6, Fast MIPv6, HMIPv6, PMIPv6, and so on. HURRICANE requires new capabilities and technologies on wired and wireless nodes. Hence, the seamless adoption of HURRICANE is challenging.

2.3.13 MobileNAT

MobileNAT [94] supports micro mobility and macro mobility across heterogeneous networks. MobileNAT deploys network address translation (NAT) devices that translate internal private address of a mobile node to an external globally unique IP address and vice-versa. MobileNAT offers following benefits: the use of private addressing in large public networks, use of heterogeneous (IPv4/IPv6) addressing schemes, flexibility in frequent changes in addressing, and easy policy enforcement. MobileNAT architecture introduces an anchor node (AN), a mobility manager (MM), and a Dynamic Host Configuration Protocol (DHCP) server and relays.

- Anchor Node: A gateway router with network address translation (NAT) support or a separate NAT device connected to a traditional router.
- Mobility manager : A mobility manager signals mobility events to the anchor node. A mobility manager may be co-located with the DHCP server.
- DHCP server and relays: DHCP assigns A_v (host identification) and A_p (physical address for routing) IP addresses to a mobile node when the mobile node moves to a new subnet. Each subnet has a DHCP relay that forwards the DHCP requests to DHCP servers at each domain. These DHCP relays are either co-located with the router or separate in the network.

MobileNAT also provides a thin software layer, called shim-layer, between IP layer and network interface driver in the client machine to maintain the translation rules as stated in Table 2.1.

Identity Management in MobileNAT

MobileNAT aims to solve the difficulties come with using IP address to identify a mobile node and to route data packets to the mobile node. MobileNAT provides use of two addresses for mobile node and tunneled packet forwarding.

- *Use of Two Addresses:* A mobile node has A_v (virtual IP for host identification) and A_p(physical IP for routing) addresses. An anchor node does address translation for source and destination addresses. A correspondent node sends packets with destination address, A_v. The anchor node translates A_v to A_p. Therefore, NAT is transparent to the correspondent node. The mobile node addressing is subject to two policies with different address combinations of A_v and A_p as stated in Table 2.1. There are four different address combinations. There are two policies can be applied by the anchor node: Policy1: If possible, expose A_v external to domain. Policy2: Never expose the mobile node's A_v.

- *Tunneling:* IP tunneling is used to forward the packets from a mobile node to an anchor node. The outer IP header has source address A_p, whereas inner IP header has A_v. The anchor node strips off the outer header before forwarding the packet to the correspondent node. Advantage of the scheme is less processing overhead while disadvantage is additional header overhead and increased packet size.

One of the issues that MobileNAT needs to address is how stateless autoconfiguration of IPv6 addresses will be supported.

Handoff Management in MobileNAT

When a mobile node changes domain, A_p is replaced with a new one, however, A_v stays the same. Home-NAT forwards the packets to the Visited-NAT's external address. Visited NAT transmits packets to the mobile node's new physical address. MobileNAT may apply address translation and tunneling (between Home-NAT and Visited-NAT) for inter-domain movement of the mobile node. The traffic from Visited-NAT to the correspondent node can be either (i) direct, in which case the Visited-NAT fakes its source address as that of the Home-NAT, or (ii) it can always reserve proxies through the Home-NAT (via tunnel or translation method). When the mobile node moves in a domain, Destination-NAT rules at the anchor node and the mobile node are appropriately altered for the new A_p. When the mobile node enters to a new NAT, it gets new A_v but keeps the old one until all old sessions are closed.

Table 2.1: Address translation by MobileNAT device.

	A_p	A_v	Policy 1	Policy 2
Case 1	Private	Private	$A_p \rightarrow A_{AN}$	$A_p \rightarrow A_{AN}$
Case 2	Private	Public	$A_p \rightarrow A_v$	$A_p \rightarrow A_{AN}$
Case 3	Public	Private	$A_p \rightarrow A_{AN}$	$A_p \rightarrow A_{AN}$
Case 4	Public	Public	$A_p \rightarrow A_v$	$A_p \rightarrow A_{AN}$

2.4 Discussion of Mobility Protocols

In this section, we provide a comparative analysis of the identity and handoff management approaches adopted by the Mobile IP protocols and proposed next generation mobility solutions. Table 2.2 summarizes the Mobile IP protocols while Table 2.3, Table 2.4, and Table 2.5 present the features of next generation mobility solutions. The protocols and schemes have been compared using a unified platform which includes the attributes such as infrastructure needs, mobile address related features, and handoff management features (see Appendix A for complete list).

2.4.1 Discussion of Mobile IP Protocols

This section provides the discussion of MIPv4, MIPv6, HMIPv6, and PMIPv6 focusing on their advantages and disadvantages in providing identity and handoff management. Table 2.2 provides a comparison chart presenting the features of each Mobile IP protocol.

Deploying macro-mobility protocols i.e. MIPv4 and MIPv6 for mobile node's movement within a single autonomous system domain may bring signaling overhead and handoff latency as the binding update messages use up wireless bandwidth and home network or correspondent node might be far away from mobile node's current network [95]. However, when micro-mobility protocols i.e. HMIPv6 and PMIPv6 are deployed for mobile node's intra-AS movement, they bring lesser signaling overhead and lower handoff latency because the routing tables at the home agent do not change as mobile node's address does not change within the domain. [39, 96–98].

HMIPv6, introduces scalability and also incorporates hierarchy in the flat IP structure for efficient handoff management. PMIPv6 does not require any software installation on mobile nodes because it follows a network-based handoff management approach. It can also work without depending on any existing macro-mobility protocol. Therefore, PMIPv6 accommodates changing technology and market requirements better. This characteristic of the protocol can accelerate the deployment of PMIPv6 [38]. In PMIPv6, mobile

Table 2.2: Qualitative comparison of MIPv4, MIPv6, FMIPv6, HMIPv6, and PMIPv6 ‡

Category	MIPv4	MIPv6	FMIPv6	HMIPv6	PMIPv6
Mobility Scope	Global	Global	Global/Local	Local	Local
Mobility Management	Host-based	Host-based	Host-based	Host-based	Network-based
Network Architecture	Flat	Flat	Flat	Hierarchical	Hierarchical
Target Network	IP	IP	IP	IP	IP
Operating Layer	L3	L3	L3	L3	L2 & L3
Required Infrastructure	HA	HA	AR & HA	AR & MAP	MAG & LMA
Mobility Protocol	MIPv4	MIPv6	FMIPv6	HMIPv6	PMIPv6
MN Modification	Yes	Yes	Yes	Yes	No
MN Address	HoA	HoA	HoA	RCoA & LCoA	CoA
Address Type	IPv4	IPv6	IPv6	IPv6	IPv6
Address Length (bits)	32	128	128	128	128
Address Change	Yes	Yes	Yes	Yes	No
Address Assigned by	HA	HA	HA	MAP	LMA
Tunneling	Inter-AS	Inter-AS	Intra-AS	Intra-AS	Intra-AS

‡For acronyms, please refer to the sections covering the current and future mobility solutions.

node uses less computing resources, encounters less wireless channel access delay and less wireless transmission delay as handoff is managed between mobility access gateway and the local mobility anchor [99, 100].

2.4.2 Discussion of Next Generation Mobility Solutions

In this section, we discuss the identity and handoff management methodologies of the next generation mobility solutions targeting current IP networks, cellular networks, mesh networks or a combination. They aim to provide *all the time connectivity* regardless of the networking technology.

The most popular identity management approach is the ID/locator separation approach where a mobile node's physical location and actual ID are different. Physical locator of the mobile node can be an IP address (DAIDALOS, i3, and Hi3), or routing locator resembling the router that mobile node is connected to (LISP and MILSA). The mobile node, the Internet content or a mobile user can be identified with human-readable string (MobilityFirst and XIA) or with specially encrypted ID-tags (HIP). Having content or mobile users being identified separately from the routing layer help content-aware communication and user-centric mobility. However, mapping and naming servers have to be located in the network and the communication of other nodes with these servers have to be regulated via protocols to have easy, fast, reliable and secure communication.

Table 2.3: Qualitative comparison of MobilityFirst, XIA, VMD, and Ambient Networks ‡

Category	MobilityFirst	XIA	VMD	Ambient Networks
Mobility Scope	Global	Global	Global	Global/Local
Mobility Management	Network-based	Network-based	Network-based	Host-based
Network Architecture	Flat	Flat	Tiered	Flat
Target Network	IP	IP	FCT	IP & Cellular
Operating Layer	L3 & L4	Tapa	L2 & L3	L1 - L5
Required Infrastructure	GDTN routers	XIP routers	VMD	Interfaces & components
Mobility Protocol	Not specified	Not specified	VMD	MIP, HIP, & SIP
MN Modification	Yes	Yes	Yes	No
MN Address	GUID	XID	Tiered	Flow & Bearer
Address Type	String	String/IP	Numeric	Abstraction
Address Length (bits)	Dynamic	160	Dynamic	Varies
Address Change	No	No	No	Yes
Address Assigned by	Naming service	Naming service	VMD	Interfaces
Tunneling	No	No	No	*

* Depends on the mobility protocol used by the architecture.

Table 2.4: Qualitative comparison of AKARI, i3, Hi3, LISP, and MILSA‡

Category	AKARI	i3	Hi3	LISP	MILSA
Mobility Scope	Global/Local	Global/Local	Global/Local	Global/Local	Global/Local
Mobility Management	Host-based	Host-based	Network-based	Network-based	Host-based
Network Architecture	Hierarchical	Flat	Flat	Flat	Hierarchical
Target Network	IP & Cellular	IP	IP	IP	IP
Operating Layer	ID	L3 overlay	HIP & overlay	Man-encap	HUI
Required Infrastructure	IMS & NMS	i3 servers	Rendezvous	New network	RZB & Zone
Mobility Protocol	MIP	ROAM	Hi3	MIP	MIP
MN Modification	Yes	Yes	Yes	No	Yes
MN Address	ID, name & locator	ID & IP	HIT, ID & IP	EID & RLOC	HUI & locator
Address Type	String/bitstream	IPv4/IPv6	IPv4/IPv6	IPv4/IPv6	string & IP
Address Length (bits)	Varies	32 & 128	32 & 128	128	varies
Address Change	Yes	Yes	Yes	Yes	Yes
Address Assigned by	IMS	i3	Rendezvous	New network	RZB & Zone
Tunneling	Intra-AS	No	Inter-AS	Inter-AS	No

* Depends on the mobility protocol used by the architecture.

‡For acronyms, please refer to the sections covering the current and future mobility solutions.

Table 2.5: Qualitative comparison of DAIDALOS, HIP, CARMEN, HURRICANE, and MobileNAT‡

Category	DAIDALOS	HIP	CARMEN	HURRICANE	MobileNAT
Mobility Scope	Global/Local	Global/Local	Global/Local	Global/Local	Global/Local
Mobility Management	*	Host-based	Network-based	Host-based	Network-based
Network Architecture	Flat/Hierarchical	Flat [1]	Hierarchical	Hierarchical	Flat
Target Network	IP	IP	Mesh	Mesh	Mesh
Operating Layer	L2 - L5	HIP	L2 & L3	L2 & L3	L3 & Shim
Required Infrastructure	ID manager & broker	RVS	CAP & CGW	CIC & HM	AN & MM
Mobility Protocol	(H/P)MIPv6 & HIP	MIP	CARMEN	(H/P)MIPv6	IPv6
MN Modification	No	Yes	Yes	Yes	Yes
MN Address	VID	HIT & IP	IDs	IP	IP
Address Type	IPv6	IPv4/IPv6	IPv6	IPv4/IPv6	IPv6
Address Length (bits)	128	32/128	48 & 128	32 & 128	128
Address Change	Yes	Yes	Yes	*	Yes
Address Assigned by	ID-manager	RVS	CAP & CGW	*	AN
Tunneling	*	Inter/Intra-AS	No	*	Optional

* Depends on the mobility protocol used by the architecture.

[1] [67] proposes hierarchical HIP.

‡For acronyms, please refer to the sections covering the current and future mobility solutions.

In handoff management, the popular mechanism is providing an integration framework for different network technologies and handoff management protocols to operate together (Ambient Networks, DAIDALOS, and AKARI). Integration frameworks could be very feasible for next generation networking if they are modular and extensible in allowing adaptation of new unforeseen mobility solutions. In another approach, gateways are provided to hide the technological differences between the access networks (CARMEN and HURRICANE). Each access network connects to the Internet via the gateways and handoff related messaging is generated between gateways and other wired mobility agents. Another approach in handoff management is deploying an overlay infrastructure (HIP, i3, Hi3, and LISP-ALT). This approach requires a strong tie between the identity scheme and the handoff management protocol.

2.5 Mobility Modelling-Related Works

Mobile device users change locations with different speeds by walking or using vehicles, individually or within a group. Understanding the characteristics of user mobility is important for creating systems with better mobility support. This section aims to provide literature review on mobility models. In the literature, scholars study several mobility models. The Gaus-Markov mobility model, random-walk mobility model, and their vari-

ations are used to conceptualize a mobile user as an individual entity. The cases where a mobile user's movement decision depends on other people are explained with mobility models, such as nomadic community, pursue, or reference-point group-mobility models. Furthermore, mobile users' movements on a mapped area are also studied by mobility models like freeway and Manhattan grid. In the following, we will present these mobility models. The extensive literature review of the mobility models can be found in [101–103].

The random-walk mobility model represents a mobile user that roams in random directions and speeds [102]. The random-walk mobility model has several variations, such as the random-waypoint mobility model that represents a user who pauses before changing speeds or directions. The random-direction mobility model, however, forces a mobile user to go to the edges of the simulation area before changing the speed and the direction. The probabilistic random-walk mobility model leverages a set of probabilities to determine the next position of a mobile user. In random-mobility models, the velocity of a mobile user at a current state is independent of the previous state. Furthermore, a mobile user can move freely in the simulation area, independent of others. Random-mobility models are more appropriate for pedestrian movements in which mobility is generally confined to a limited geographical area such as residential and business buildings [104].

The Gauss-Markov mobility model [105] introduces randomness in user's movement direction and speed gradually. At the initial state, a mobile user is assigned a speed and a direction. Later, at fixed time intervals, the speed and the direction of the mobile user is updated according to the previous state and random variable.

The mobility of a user on a mapped area, such as on highways or streets, is formulated under mobility models, such as freeway and Manhattan-grid mobility models. The freeway mobility model simulates a mobile user's movement on highways [106]. In this model, a mobile user is restricted to its lane on the freeway. His speed may depend on his or the preceding user's speed. The Manhattan-grid mobility model [107] simulates a mobile user's movement on streets. Therefore, the simulation area is divided into squared blocks and users are modeled as pedestrians moving on the vertices of the squares (streets). Initially, the nodes are randomly distributed on the streets. Each user chooses a direction and a velocity. If a user reaches a corner, the user changes direction with a certain probability. The velocity of the user is changed over time.

The fluid-flow mobility model [108] represents a flow of mobility traffic as the flow of a fluid. The model formulates the amount of traffic flowing out of a region to be proportional to the population density within the region, the length of the region boundary, and the average velocity. The fluid flow model is more suitable for users with high mobility, infrequent speed, and direction changes [109, 110]. With the assumption of a uniform

distribution of both the mobile users' position and the movement direction, the boundary crossing rate can be approximated by the use of fluid-flow mobility model [104, 109].

Tracy et al. [101] present the group mobility models where each individual is affected from group behavior. For example, the nomadic-community mobility model represents the mobile users move together from one location to another. In this model, individuals maintain their own personal spaces by moving in random directions within the flow. The pursue mobility model formulates the movement of a group that follows a given target. Reference-point group mobility model studies random motion of a group of people as well as random motion of individuals within a group. Group movements are based on the paths traveled by a logical center for the group. Individual mobile users randomly move around their own predefined reference points, which is affected by the movement of the logical center of the group.

Researchers have also derived mobility patterns from real network data. Afanasyev et al. [111] study traffic, mobility, and data usage patterns of mobile users in Google WiFi local area networks in Mountain View, CA. Other similar studies and real-world wireless local area network data can be found in [112].

In Chapter 4, we study the performance of our proposed architecture in comparison with IPv6-based mobility protocols. In our performance studies, the mobile user travels between access points, following the trajectory shown by the red line, (in Fig. 4.1) at a steady speed. The mobile user waits at each destination for a specific amount of time before starting to move toward the next access point. This path is chosen because, along the entire path, the mobile user performs two different types of handoffs that we aim to observe, which are intra-cloud and inter-cloud handoffs. These handoffs enable us to assess the effect of architectural differences, such as mobility agent locations, and the fundamentally different ways of handling handoffs, such as network-based and host-based handoffs. We are interested in how mobility protocols react after the mobile user establishes a connection to an access point. Therefore, we structure the movement of the mobile user on a trajectory with the steady speed, and we do not use different mobility patterns. As the focus is on handoff-performance, we limit the investigations and results to a single user that goes through the handoff process.

In Chapter 6, we conduct a mobility study aiming to retrieve the number of handoffs that a mobile user makes. Further, we also provide a method to find the location of the mobility agents in the architecture that manage the user's handoffs. We propose a mobility model that represents a user who spends most of his time on specific locations, such as home, school, work, and the mall. The user rarely goes to different locations at great distances. In our analysis, we consider the repeatedly visited places as the center of

the roaming area.

We introduce a new mobility model because the aforementioned mobility models do not consider this centered-movement characteristic of a user. For example, random-walk mobility models focus on the randomness in the direction and the speed of the user's movement without considering where the user moves more frequently. In our proposed mobility model, we are interested in the user's overall roaming with average speed instead of his individual steps. The fluid-flow mobility model is used extensively in finding the boundary crossing rate of a group of users. However, in that mobility model, the user is modelled as going out of a region all the time, proportional to the population density in a given area. In our proposed mobility model, a mobile user has a tendency to move in the center of his roaming area. We further do not consider the group-based mobility models because we assume a user gives his movement decisions independently from others. We study an individual user who may have unique roaming activities not affected by group motion. We do not consider the effect of buildings, highways, or streets on the mobile user, but it is considered by the mapped-based mobility models, such as freeway or Manhattan-grid mobility models. For this reason, our proposed architecture is built on the already-existing tiered structure between ISPs and ASes in the Internet. Further, we assume the architecture deployment is not limited to residential or business buildings. We, instead, study the mobility of a user focusing on his number of handoffs and the location of his mobility agents on the architecture in a given period of time, roaming range, and speed of movement. We locate the roaming range of the mobile user randomly within the area that is covered by the proposed mobility architecture. Then, given the user's mobility characteristics, we retrieve the location of the mobility agent that will handle the mobility.

In the last two paragraphs, we present our motivation behind the use of the specific mobility model. However, this model is used just to illustrate a methodology on finding the number of handoffs and the location of the related mobility agents that will handle the mobility. Users can have different mobility characteristics and our aim in Chapter 6 is to show that our architecture and handoff cost optimization is user-centric. In the aim of reaching that goal, user handoff information is very important factor and hence we will use it in the handoff cost optimization equation that is presented in the same chapter.

2.6 Handoff Cost Optimization-Related Works

This section presents research efforts toward the improvement of network resource usage and quality-of-service provisions during a user mobility support in cellular and IP networks. Existing optimization studies mostly focus on different mobility parameters to minimize the handoff cost that can be caused by latency, signaling overhead, etc. A

few of the studies, however, aim to improve the handoff sub-processes to decrease the handoff cost.

Lin proposes a location-tracking strategy to reduce the location update cost in personal communication service (PCS) networks [113]. In that study, users' movement patterns and call traffic significantly impact the location-tracking operations. The proposed algorithm performs well for the cases where the call-to-mobility ratio is low, the registration cost is large, and the user presents various residence times.

Akyildiz et al. [104] examine location-tracking costs, call-loss rates, and paging delays in cellular networks. They introduce the concept of a boundary location area that is determined according to the mobile users speed and the network load. An inter-system paging system is created by the proposed boundary location register concept. These methods help to reduce call-loss rates, paging delays, and signaling costs due to location tracking.

Xie et al. [114] propose a dynamic regional-location management scheme for mobile IP networks to decrease the number of messages sent to home networks for location updates and to decrease the signaling costs due to data packet delivery. In that manner, they provide a framework to find an optimal regional network size that considers user mobility, packet arrival patterns, and network density. The aim is to minimize the network resource usage of a mobile node.

The aforementioned cellular network studies focus on the extra messaging that has to be done to track the location of a mobile user and update the record of the related network devices. These processes mainly consume the network resources by occupying the bandwidth and using the devices' processing powers. These costs primarily affect the service providers, and hence, these optimization studies are conducted from a service-provider perspective.

The signaling cost incurred in MIPv6, Fast MIPv6, HMIPv6, and PMIPv6 is analyzed in [42]. Handoff-management-related messaging, packet delivery, and packet tunneling costs are formulated to find the impact of the sub-processes on the network resource consumption. They also conduct a comparative performance study in terms of the signaling cost of the aforementioned protocols. Lee et al. [110] examine the wireless power consumption cost of HMIPv6 and PMIPv6 due to location update and packet delivery. The aforementioned attributes are mainly affecting the network resource usage, and therefore, the optimization of these costs is the main concern of service providers. At every handoff, the service provider has to update its routers and mobility agents, and transfer data packets to the user's new location as part of its handoff support.

Mobile users may observe a delay or QoS degradation in their connectivity due to the handoff, rather than the signaling overhead on the network. Kong et al. [39] examine the aforementioned IPv6-based protocols in terms of delay caused during a handoff. The goal of this study is to find the tasks causing delay in handoff and then focusing on the best protocol that brings the lowest handoff latency in the interest of the mobile user.

Vilhar et al. [115] study different network topologies, such as tree topologies, and they examine the success of mobile anchor-point-selection algorithms in minimizing the location update and packet delivery cost in HMIPv6. Pack et al. [109] study the optimal hierarchy level and network size for HMIPv6 that minimize the location update and packet delivery costs. These two studies aim to find the optimum network topology to provide a better handoff performance in terms of network resource usage.

Jeon et al. [116] propose a fast handoff mechanism to decrease the handoff delay in PMIPv6. Further, the route between the mobility agents is optimized to provide a low packet-delivery cost. A reactive handoff scheme for IPv6-based mobile networks is introduced in [117] to reduce the handoff delay by optimizing the movement detection and address configuration processes that happen during the handoff. Dutta et al. [118] propose a media-independent, pre-authentication-based handoff scheme for IP-based mobile networks by obtaining the network parameters before the handoff occursleveraging the predictive handoff information. This proactive handoff mechanism provides a low handoff delay and fewer data packet losses. The presented reactive and proactive handoff methods are changing the way that a handoff is initiated to decrease the handoff delay, which impacts a user's mobility experience.

PERIMETER [119] provides a Quality of Experience (QoE) model to improve the users' perception of mobility. The proposed model considers QoE signaling, user preferences, network condition, and content adaptation support during the network selection process. Calvagna et al. [120] propose a model to balance the overall cost of vertical handoffs between GPRS and WLAN. The proposed scheme introduces a new vertical handoff policy that considers not only network characteristics but also transport and application layer preferences to find the best network aligned with a user's actual needs. Islam et al. [121] present a user-centric service provisioning (such as service subscription, session setup, profile management, and privacy) for IP multimedia sub-systems that benefit not only users but also service providers and network operators with new business models.

2.7 Summary

Mobile Internet usage continues to increase at a dramatic pace. In this chapter, we provide the previous research works in the literature related to the subjects studied in the next chapters. We provide a literature review of mobility models and handoff-cost-optimization studies. Then, a survey of identity and handoff management solutions for the next-generation mobility architectures and protocols is presented. In addition, a qualitative comparison of mobile IP protocols and next-generation mobility protocols on a uniform platform is provided. The goal of this survey is to provide an unbiased report on all solutionscurrent and future. Along with the several approaches adopted to overcome the current challenges in identity and handoff management, the future Internet is envisioned to be open and receptive to different approaches based on future needs. The chosen approach will be dependent upon several factors, such as user mobility characteristics, type of applications, QoS provisions, handoff cost, etc. Our proposed identity- and handoff-management approaches are presented in the next chapter.

Chapter 3

Virtual Mobility Domain (VMD) Architecture

The population of mobile users seeking connectivity to the Internet has been growing over the years, spurred by the capabilities of handsets and the increasingly rich Internet content and services. Mobility management to enable efficient Internet access for users on the move is thus gaining significance. With future Internet design initiatives gaining momentum, it is important that these initiatives consider mobility management as an integral part of the design. In the previous chapters, definition of user mobility management, its challenging aspects, and ongoing research efforts to provide better user mobility experiences are presented in detail. Stating the challenges and our motivation in the preceding chapters, the questions that we aim to answer with our research are as follows:

- *Identity management:* Can we develop an addressing scheme that is more mobility friendly? An ideal mobility-friendly addressing scheme would be powerful in identifying and tracing the mobile node. The scheme should be capable of supporting

* Portions of this chapter previously appeared as:
H. Tuncer, A. Kwasinski, and N. Shenoy, Performance Analysis of Virtual Mobility Domain Scheme vs. IPv6 Mobility Protocols, *Elsevier Computer Networks Journal*, Volume 57, Issue 13, 9 September 2013, Pages 2578-2596, ISSN 1389-1286.
H. Tuncer, Y. Nozaki, and N. Shenoy, Virtual domains for seamless user mobility, in *Proceedings of the 9th ACM international symposium on Mobility management and wireless access, ser. MobiWac11*. Miami, FL, USA, 2011.
H. Tuncer, Y. Nozaki, and N. Shenoy, Virtual Mobility Domains - a mobility architecture for the future internet, *Communications (ICC), 2012 IEEE International Conference on*, vol. 2774, no. 2779, pp. 10-15, June 2012.
H. Tuncer, Y. Nozaki, and N. Shenoy, Seamless user mobility in Virtual Mobility Domains for the future internet, *Communications (ICC), 2012 IEEE International Conference on*, vol. 5860, no. 5865, pp. 10-15, June 2012.

the changes in a mobile node's physical connection point. Furthermore, the address length should not be so long that it consumes costly wireless bandwidth. The provided addressing space should be large enough to support the growing number of mobile nodes in future wireless networks.

- *Handoff management:* Is there a more efficient handoff-management scheme? The handoff-management process is executed when a mobile node changes its subnet or domain. In that case, the previous network, the new network, and the correspondent nodes that the mobile node is in communication with may need to be informed about the movement of the mobile node in order to allow the previous session to continue. Eventually, the routing devices' routing table entries need to be changed. The efficiency of the mobility protocol relies on a minimum number of changes on wired and wireless devices. The successful implementation of such mobility protocol will lead to low latency and less signaling load on the network, and therefore there will be less computational resource usage, eventually creating seamless handoff management.

This dissertation discusses the design and implementation of a novel, future Internet-mobility architecture called Virtual Mobility Domain (VMD) with the aim of answering the questions above. The VMD proposes a novel addressing scheme with a unique address-acquisition mechanism and a seamless intra-AS and inter-AS handoff-support mechanism. The VMD is built to work on the FCT internetworking model that is proposed for a future Internet. The FCT internetworking model introduces a tiered structure, a tiered addressing model, and a unique packet-forwarding scheme. The VMD supports network-based mobility management and leverages the tiered structure to provide collaborative handoff management in a domain that can span several networks. A novel, collaborative mobility-management scheme offers dynamic handoff management with respect to user mobility patterns and, hence, is expected to bring less signaling overhead and lower latency. The VMD is user centric because it is possible to overlap the VMDs of different scopes and allow a user to select a VMD that is most suited to his roaming needs. The proposed mobility architecture is distinct from others by not using IP addressing and classic routing protocols and by deploying user-centric overlapping mobility domains.

The rest of this chapter is organized as follows. Section 3.1 presents the fundamentals of the FCT internetworking model. The VMD mobility architecture, with its main attributes, is presented in Section 3.2. Deployment of the VMD in an AS, and the mobility support in an AS, are presented in Section 3.3. Then, the VMD deployment across multiple ASes and ISPs to support macro-mobility is presented in Section 3.4. A summary of the architecture is given in Section 3.5.

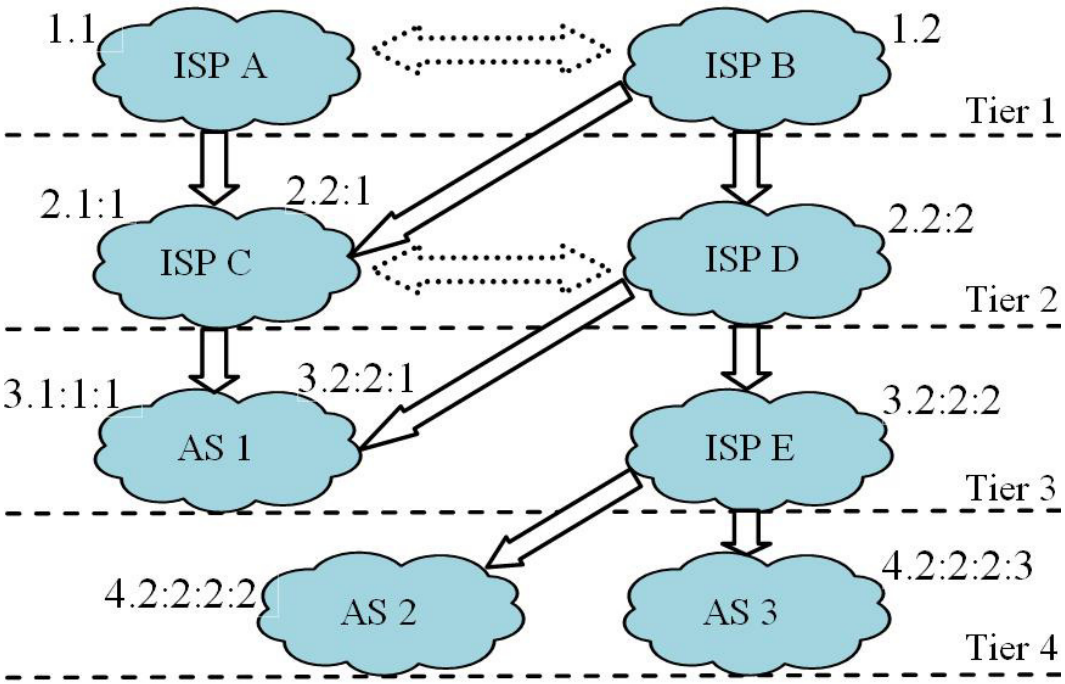


Figure 3.1: The FCT model is applied to ISP networks and ASes. Arrows show the provider-customer relationships. Dotted arrow shows peering relationships. Each arrow represents a single link or multiple links [2].

3.1 The Floating Cloud Tiered (FCT) Internetworking Model

The new tiered Internet architecture leverages the tiered structure that exists among ISPs to define their business relationships. However, to overcome topological rigidity in the tiered ISP structure and to enable easy attachment of entities such as networks or ASes across the tiers, granularity and modularity were introduced. Granularity was introduced through the concept of network clouds [3] where a network cloud can be an ISP network, a point of presence in an ISP, an AS, or a set of routers. Modularity was introduced by decoupling the relationships between the network clouds through the use of a nesting concept. Thus, the model was named the Floating Cloud Tiered (FCT) internetworking model. The FCT internetworking model allows network clouds to connect at any tier without impacting their internal address or structure when nesting is adopted [2].

The goal of the FCT is to overlay the tiered structure (existing among ISPs) on the meshed Internet topology to facilitate efficient routing using the tiered addresses. Inherently, there is a hierarchy in the tiered structure which is being used in the VMD. To explain

VMD in the context of the FCT internetworking model, we provide a description of the FCT model and its operation using ISP networks in Fig. 3.1. Each ISP or AS is identified with a tiered cloud address, which is a function of the tier in which the network cloud resides and its association to its service provider network clouds. In Fig. 3.1, ISP A, which is the first cloud in tier 1 has a cloud address 1.1 following a format *TierValue.MyCloudID*. ISP B similarly has a cloud address 1.2. ISP C is in tier 2 and connected to both ISP A (via cloud address 2.1:1) and ISP B (via cloud address 2.2:1) simultaneously. The cloud address format in this case is *TierValue.ParentCloudID:MyCloudID* where the ParentCloudID is inherited from the upper tier clouds. ISP D is connected to ISP B through cloud address 2.2:2. Similar provider-rooted address can be noticed in [122, 123]. In Fig. 3.1, each arrow represents a single link or multiple links. Each link may get different address or same address. When they are assigned same address, the address and each link's port information needs to be shared among the routers in the same cloud [2].

The packet forwarding decision across the clouds (up, down, or sideways) depends on the relative positions of the source and destination clouds. To illustrate, if AS 3 (4.2:2:2:3), source cloud, wants to send packet to AS 1 (3.2:2:1), destination cloud, then source compares its address with the destination's address to determine the tier of a common parent (or grandparent) cloud. In this case, common parent cloud will be ISP D at tier '2'. The remaining fields in the destination address (after the common part) are then appended to the TierValue to provide the forwarding address. In this case forwarding address will be 2.1. All the intermediate clouds between AS 3 and ISP D forward the packet upwards, using the tier value. When the packet reaches ISP D, then it identifies that the destination is at tier 3 because of the one address field following the tier value. It replaces the TierValue with 3 and forwards the packet down to the destination cloud. However, if there were a peering link between ISP E and AS 1, the border routers in ISP E would have routing table entries for the sibling cloud connections and ISP E could forward the packet directly to the AS 1. The FCT architecture, its tiered addressing scheme, support of cloud movement, the packet forwarding and the study of Internet's topology are explained in more detail in [2, 124].

The tiered structure applied to ISP networks in the above example can be applied within an AS or a network. If such were the case as illustrated in Fig. 3.2, the border or backbone routers can be associated to a network cloud in tier 1, the distribution routers can be associated to a network cloud in tier 2. The access routers and subnets can be associated to tier 3. In the next section, the VMD architecture is presented by applying the tiered structure to the AS network in Fig. 3.2.

The tiered addresses proposed for the FCT internetworking model is used to forward packets to the mobile node. In VMD, the relationship between network clouds in the

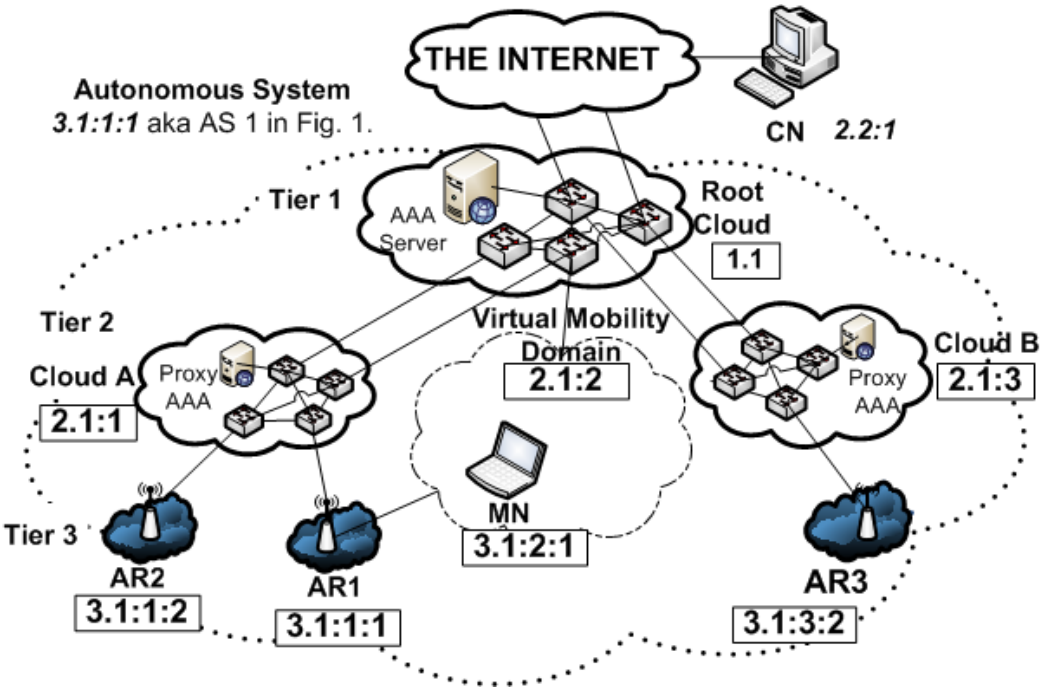


Figure 3.2: VMD is applied to an AS. The backbone routers form the cloud in tier 1. Cloud A and Cloud B in tier 2 represent network clouds that have distribution routers. Access routers reside in tier 3.

tiered structure and the inheritance information in the tiered address allow a mobile node to roam within a defined VMD using a single address.

3.2 Virtual Mobility Domain

The network cloud concept defined under the FCT internetworking model can be extended to *virtual network clouds*, where the set of devices in the network cloud are not geographically or physically constrained to a locality. A virtual network cloud thus defines the boundary of a VMD. A VMD has to be supported by an upper tier cloud, thereby allowing mobile nodes that have an address in the VMD to roam within the scope of that upper tier cloud i.e. within all network clouds that are connected under the upper tier cloud. A mobile node's movement in the VMD is *collaboratively managed* by the upper tier cloud and all network clouds under the upper tier cloud. The details are explained in Sections 3.2 and 3.2.

A VMD can be applied to any tier or cloud to cover single AS, several ASes or ISPs in

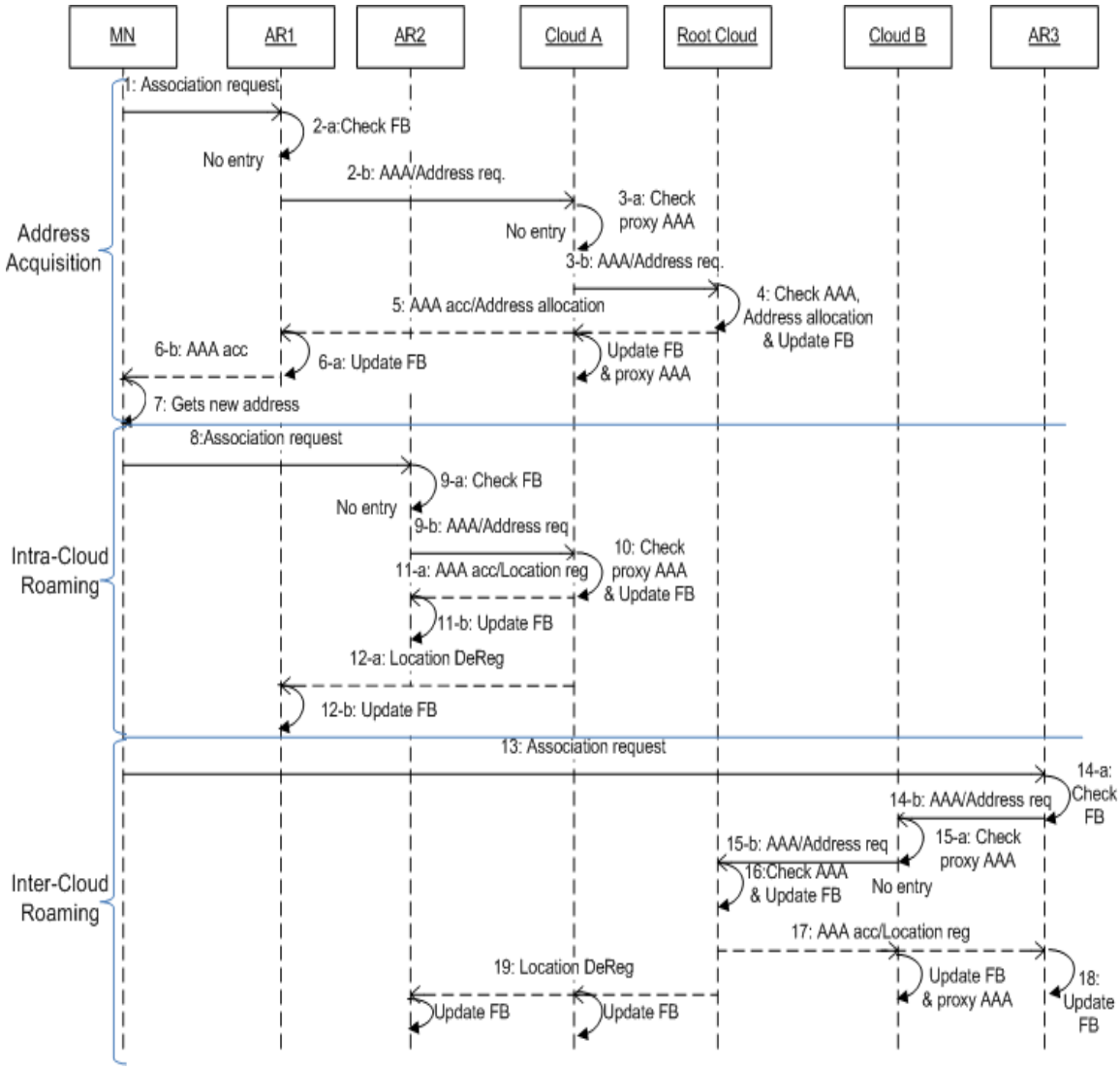


Figure 3.3: Control message flow for intra-AS roaming in the VMD.

the current Internet. The agreements across the clouds are necessary in extending VMD across ASes or ISPs [3]. In this section, we focus only on VMD deployment within a single AS. VMD deployment across multiple ASes and ISPs can be found in Section 3.4.

VMD Implementation

Fig. 3.2 is used to explain the VMD deployment in an AS, such as AS 1 (Fig. 3.1) with a cloud address 3.1:1:1. The network clouds and tiers concept have been applied within the AS. Nesting of tiers as described in [2] allows the tier numbers within the AS to start at 1.

The backbone routers and an AAA server form the network cloud in tier 1, that is named as Root Cloud. Cloud A and Cloud B are the tier-2 network clouds comprising of distribution routers that connect to the backbone routers in Root Cloud in tier 1. Each tier-2 cloud has *a proxy AAA server* to maintain mobile node profile while the mobile node is roaming under the same cloud. The access networks and devices are in tier 3. Each cloud maintains a *forwarding base* (FB) to track the physical location of the mobile node roaming within the cloud. Cloud addresses are noted beside the clouds and the devices.

In this section, we assume that there is only one Root Cloud. In the case that there were two Root Clouds that covered all the clouds at lower tiers in an AS, any one could serve to provide the coverage under the VMD. In Fig. 3.2, the VMD cloud is defined under the Root Cloud and has a cloud address 2.1:2. Mobile nodes supported under this domain will thus get an address in tier 3, namely 3.1:2:n, (where 'n' can be any integer value). Mobile nodes can roam in any of the network clouds under Root Cloud namely Clouds A and B. The mobile node in Fig. 3.2 has an address 3.1:2:1. The global address of the mobile node will however be 3.1:1:1{3.1:2:1} where the mobile node's local VMD address is appended to AS cloud address following the nested notation defined in [2]. Forwarding data packets to the mobile node within the VMD however will use the internal address 3.1:2:1 and is described in Section 3.3.4.

Collaborative Mobility Management

The AAA server located in Root Cloud maintains the profiles and authentication details for mobile nodes roaming within the AS. When a mobile node is registered to a VMD, a proxy AAA server in a tier-2 cloud makes a copy of the mobile node profile from the AAA server in tier 1. The mobile node may roam in the vicinity of AR1, AR2 or AR3, and its physical location is tracked by the forwarding bases maintained at the clouds. For example, when a mobile node is connected to AR1 or AR2, the distribution routers in Cloud A will maintain an entry in the forwarding base for the mobile node that points to the mobile node's current access router, while the backbone routers in tier 1 will have a forwarding base entry which indicates that the mobile node is physically located under Cloud A. If the mobile node moves between access routers i.e. AR1 and AR2 under the same tier-2 cloud, then Cloud A will be the common anchor cloud (CAC) for AR1 and AR2. We call the entity that manages mobile node mobility and handoffs as mobility agent

(MA). The mobility agent in the common anchor cloud manages the mobile node mobility with the help of proxy AAA servers and forwarding bases. Hence, no communication with Root Cloud is required.

3.3 VMD Intra-AS Roaming Support

This section describes the processes and the signaling for intra-AS roaming using VMD. The mobile node is assumed to be in its home network, which is a valid assumption as the focus is on VMD based micro-mobility support. To roam within a VMD, a mobile node requires an address. Address acquisition in the VMD is described in Section 3.3.1. This is followed by the signaling required for the mobile node's intra-cloud roaming, e.g. from AR1 to AR2 (in Fig. 3.2), which is described in Section 3.3.2. When the mobile node roams from AR2 to AR3 (inter-cloud), the signaling now involves Root Cloud in tier 1, and is described in Section 3.3.3. The information flow and related processes during the handoff management are shown in Fig. 3.3.

3.3.1 Address Acquisition

In Fig. 3.3, steps numbered 1 to 7 describe the processes and message exchange for address acquisition by a mobile node.

1. The mobile node receives beacon packets from AR1. The mobile node decides to associate with AR1 and sends a layer 2 association request to AR1 along with its UniqueID, which can be its MAC address.

2. (a) AR1 checks its forwarding base, for an entry for the mobile node.
   (b) If there is no entry, AR1 will send an AAA request (AAA_req) message that includes the mobile node's UniqueID, and AR1's network address. AR1's address is included to receive the response back. The aim of sending AAA request message is to authenticate the mobile node and to get an address for the mobile node.

3. (a) The AAA request message is received by a router in Cloud A that is responsible of mobility management. The router checks the proxy AAA server maintained by the cloud for an entry for the mobile node.
   (b) If there is no entry, which could happen if this is the first time that the mobile node is roaming into Cloud A coverage area, the router will forward the AAA request message to Root Cloud in tier 1.

4. A router in Root Cloud which receives the AAA request message will check with the AAA server to authenticate the mobile node. An address is then allocated to

the mobile node from the VMD cloud. Let the mobile node address be 3.1:2:1. The forwarding base in Root Cloud is updated with the mobile node address and Cloud A address as a downlink address to be used to forward packets to the mobile node.

5. Root Cloud then sends an AAA acceptance (AAA_acc) message to AR1 in Cloud A. While the AAA acceptance message is delivered to AR1, the router in Cloud A that handles mobility management records AR1 address in its forwarding base as a downlink address for forwarding packets destined to the mobile node. The AAA profile of the mobile node is also maintained by a proxy AAA server in Cloud A to be used later for the mobile node's intra-cloud roaming.

6. (a) When AR1 receives the AAA acceptance message, it updates its forwarding base with the mobile node's address.
   (b) AR1 forwards the AAA acceptance message to the mobile node, with the newly allocated address for the mobile node.

7. From the AAA acceptance message, the mobile node gets its address. Subsequently, the mobile node will use this address for roaming under any cloud in the AS.

3.3.2 Intra-Cloud Roaming

In Fig. 3.3, steps numbered 8 to 12 describe the processes and message exchanges to support intra-cloud roaming of the mobile node.

8. The mobile node moves into the coverage area of AR2 and sends an association request to AR2.

9. (a) AR2 checks its forwarding base for an entry using the mobile node's UniqueID.
   (b) If there is no entry, an AAA request message will be sent to a router responsible for mobility management in Cloud A.

10. Since the mobile node has been already registered in Cloud A, an entry for the mobile node will be located at the proxy AAA server. The forwarding base entry for the mobile node is updated with AR2 address as a downlink address to forward packets to the mobile node.

11. (a) An AAA acceptance message is then sent to AR2 by the cloud router.
    (b) AR2 updates its forwarding base with the mobile node's address.

12. (a) A location deregistration (L_dereg) message will be sent to AR1 by the router in Cloud A.
    (b) AR1 removes the mobile node address from its forwarding base.

In the above transactions, mobility management related messages are limited to Cloud A in tier 2 due to the collaborative management supported by the VMD.

3.3.3 Inter-Cloud Roaming

If the mobile node moves from AR2 to AR3, located in Cloud B, steps numbered 13 to 19 of Fig. 3.3 are executed. The mobility control message exchanges are similar to the movement of the mobile node from AR1 to AR2, the difference is that the AAA request message is forwarded to Root Cloud in tier 1 by Cloud B because the mobile node is roaming in Cloud B domain for the first time and hence proxy AAA server in Cloud B does not have the mobile node profile. Root Cloud receives the AAA request, retrieves the mobile node's previously allocated address and then sends a AAA acceptance message to AR3. The forwarding base and the proxy AAA server in Cloud B and then the forwarding base at AR3 are updated with the AAA acceptance message. A location deregistration message is sent from Root Cloud to Cloud A to remove the mobile node from its forwarding base, and a similar notification is sent from Cloud A to AR2.

3.3.4 Packet Forwarding

Assume a correspondent node with address 2.2:1 as shown in Fig. 3.2 sends data packets to the mobile node using the mobile node's global address namely 3.1:1:1{3.1:2:1} where the mobile node's local VMD address is appended to AS cloud address. The data packets are delivered to the AS by the FCT packet forwarding scheme [2]. The backbone routers in Root Cloud that receive data packets will remove the AS cloud address. Within the AS cloud, data packets to the mobile node will be delivered using its local VMD address 3.1:2:1. At Root Cloud, a router will check its forwarding base for the current network cloud of the mobile node. Packets will be forwarded to that cloud, where a router will check the forwarding base to locate the access router to which the mobile node is associated and forward the data packets to that access router. The access router will then forward the packets to the mobile node. In this packet forwarding scheme, tunneling is not required avoiding the processing and overhead due to tunneling.

3.4 VMD Inter-AS Roaming Support

This section focuses on the VMD deployment across multiple ASes and ISPs, the address acquisition in the VMD, then the collaborative mobility management scheme applied for macro mobility. In Fig. 6.1, VMD 3 is deployed under ISP C. The mobile nodes connected to the VMD 3 roam within ISP C, AS 1, and AS 2. When VMD is deployed to upper tiers as VMD 1 and VMD 2 the scope of the mobility domain expands to domain 1 and 2, respectively. With these overlapping mobility domains, the proposed mobility architecture offers mobile users flexibility of registering to any mobility domain depending on their mobility patterns.

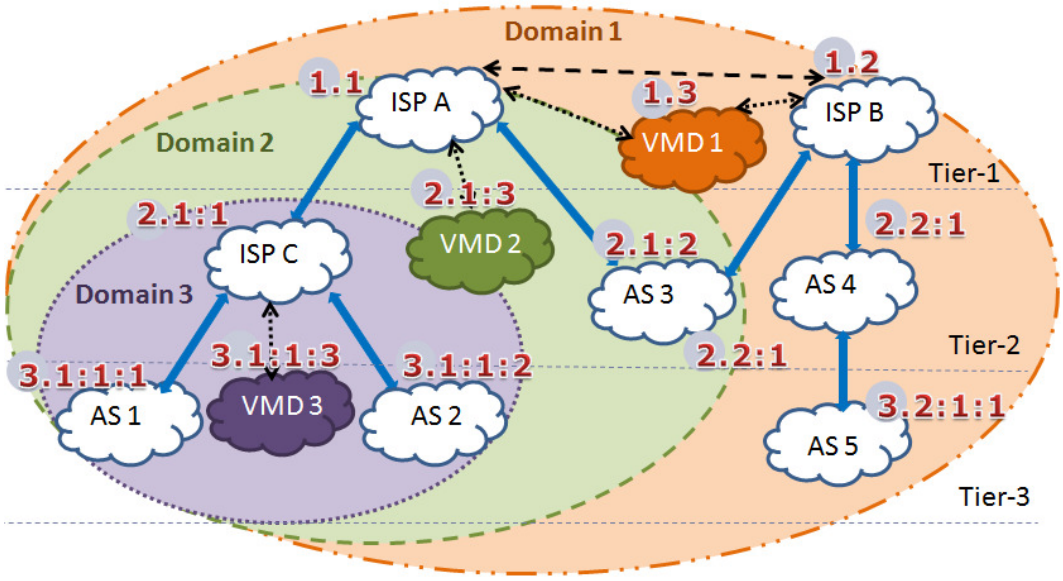


Figure 3.4: The FCT Internetworking model is applied to ISPs and ASes in the Internet. Thick and dashed arrows show the provider-customer and peering relationships, respectively. Dotted arrows show the VMDs' deployment to upper tier clouds, hence three different mobility domains are created.

Fig. 3.5 illustrates the detailed view of domain 3 in Fig. 6.1. AS 1 and AS 2 can be a university AS. In each AS, nesting described in [2] has been adopted, hence tier numbers within each AS start at 1 and nesting addresses are in rectangle in Fig. 3.5. Backbone routers and a proxy AAA server form a network cloud namely the Root Cloud in tier 1. Each college network comprised of distribution routers in the university and proxy AAA server is considered as a cloud in tier 2. The access routers and networks are in tier 3 and connected to the distribution routers in tier-2 clouds.

The VMD cloud is deployed under ISP C and resides in tier 2. The VMD cloud has a cloud address 3.1:1:3. Mobile nodes under this VMD cloud will thus get an address in tier 4, namely 4.1:1:3:n, (where 'n' can be a unique integer value). Thus, the mobile node in Fig. 3.5 has global address 4.1:1:3:1. The mobile node can get VMD service in the area that spans ISP C, AS 1, and AS 2. The mobile node's movement is *collaboratively managed* by these network clouds. To accomplish that, the clouds in the ASes have proxy AAA servers. They also have forwarding bases to track the physical location of the mobile node.

In the following sections, the address acquisition, intra-cloud, inter-cloud and inter-AS roaming support are presented in the case that VMD is deployed across ASes and ISPs

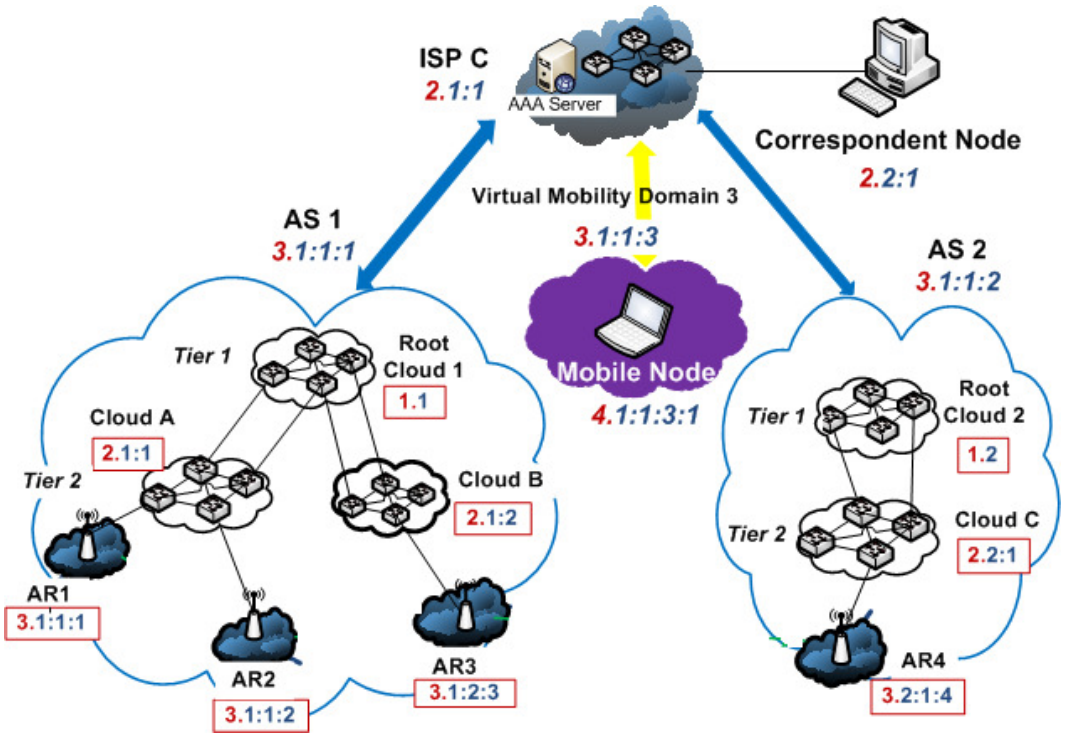


Figure 3.5: Detailed view of the domain 3 in Fig. 6.1 and it is explained in Section 3.4.

as illustrated above.

3.4.1 Address Acquisition

The Mobile node sends a layer-2 association request to an access router that it wants to connect, e.g. AR1. layer-2 association request has the mobile node's uniqueID e.g. MAC address. AR1 checks its forwarding base for an entry for the mobile node. This is the first time that the mobile node connects, hence there will be no entry at the forwarding base. AR1 sends an AAA request message that includes the mobile node's uniqueID and AR1's address. The AAA request message is received by an upper cloud, Cloud A. It checks its proxy AAA server for the mobile node profile. Again, there will not be an entry since this is the mobile node's first time in the domain. Then, the AAA request message will be forwarded to upper tier clouds till confirming the mobile node's profile. The mobile node's profile is confirmed at the cloud where the VMD is deployed, that is ISP C, assuming that the AAA server at ISP C has the mobile node profile details. 4.1:1:3:1 is then allocated to the mobile node from the VMD cloud. The forwarding base in ISP C includes AS 1 address as a downlink address to be able to forward the data packets destined to

the mobile node address. ISP C then sends an AAA acceptance message to AR1. The intermediate clouds that are Root Cloud 1 and Cloud A copy the mobile node profile into their proxy AAA servers and update their forwarding bases with proper downlink addresses for the mobile node during the message transmission. Upon receiving the AAA acceptance message, AR1 updates its forwarding base and forwards the message to the mobile node. The mobile node starts using the address retrieved from the AAA acceptance message. mobile node gets VMD service using the same address in VMD 3 coverage area that includes ISP C, AS 1 and AS 2.

3.4.2 Intra-Cloud Roaming

The mobile node's movement within the same cloud, e.g. from AR1 to AR2 is considered as an intra-cloud mobility. The mobile node starts layer-2 association to AR2 and then AR2 checks its forwarding base. If there is no entry, it sends AAA request to Cloud A. Cloud A is now the common anchor cloud between AR1 and AR2, hence it can confirm the mobile node identity via its proxy AAA server. Cloud A updates its forwarding base with AR2 address as downlink address for the mobile node. Then, it sends the AAA acceptance message to AR2. AR2 adds the mobile node address to its forwarding base and then AR2 starts accepting the mobile node's data packets. Cloud A also sends the location deregistration to the old access router i.e. AR1 to remove the mobile node's entry from its forwarding base. The collaborative mobility management scheme limits the mobility related messaging to the common anchor cloud, Cloud A via proxy AAA server deployment.

3.4.3 Inter-Cloud Roaming

The mobile node's inter-cloud movement within the AS, e.g. AR2 to AR3 is controlled by a mobility agent in the common anchor cloud, that is Root Cloud 1. AR3 sends an AAA request message to upper clouds, and Root Cloud 1 is the only cloud be able to answer the request. In response, Root Cloud 1 sends a AAA acceptance message to AR3 and a location deregistration message to the previous downlink, Cloud A. Cloud A then forwards the message to AR2. So these access routers and the intermediate clouds update their forwarding bases and proxy AAA servers.

3.4.4 Inter-AS Roaming

If the mobile node moves to another AS, e.g. from AR3 in AS 1 to AR4 in AS 2, then none of the clouds in the new AS can confirm the mobile node identity. Therefore, the AAA request message is forwarded to ISP C, which is a common anchor cloud in this case. ISP C checks its AAA server and then updates its forwarding base with AS 2 address as a downlink address for the mobile node. It sends a AAA acceptance message to AR4 and

a location deregistration message to the previous donwlink, AS 1. The clouds in the AS 1 keep forwarding the message till reaching to AR3. The aim is for these nodes to update their forwarding bases and proxy AAA servers.

3.5 Summary

We propose a VMD as a future Internet-mobility architecture that is designed to work on the FCT internetworking model. The FCT internetworking model implements a tiered Internet structure where each entity, e.g., ISP, AS, or a network, is identified with a tiered address. The FCT internetworking model also introduces a unique packet-forwarding scheme, leveraging the tiered addressing.

We designed the VMD along with the FCT model and leveraged the tiered structure and addressing introduced by the FCT model. The VMD introduces a virtual network-cloud concept that defines the boundary of a roaming domain where mobile nodes roam with a single address assigned by the VMD. The actual network clouds in the VMD handle a handoff of the mobile node in a collaborative manner with the help of forwarding bases and proxy AAA servers in each cloud. A virtual network cloud can be deployed to any network cloud in the FCT internetworking model; hence, the VMD can be extended to any domain size, e.g., single AS or across multiple ASes and ISPs. The VMD eliminates the differentiation of micro- and macro-mobility by handling intra-cloud, inter-cloud, and inter-AS handoff of a mobile node in the same manner. Therefore the VMD is user centric as it allows mobile users to register to VMDs of varying scopes, which are suited to the users roaming needs. We limit this chapter to the presentation of the design fundamentals of the VMD including various deployment scenarios, address acquisitions, and handoff supports. We will present the performance analysis of the VMD compared to MIPv6, HMIPv6, and PMIPv6 in the next chapter.

Chapter 4

Performance Analysis of the VMD

Analytical- and simulation-based performance studies of intra-AS and inter-AS handoff support of a VMD, in comparison with MIPv6, HMIPv6, and PMIPv6, are presented in this chapter. The aforementioned IPv6-based protocols are chosen specifically because they present fundamentally different ways of handling a handoff, and we aim to investigate the contributions due to locations and functions of the different mobility agents on the network. It is not possible to compare the VMD with other recent future Internet solutions because most are in the research phase. However, a comparison with MIPv6, HMIPv6, and PMIPv6 will show the relative performance improvements achieved with the VMD, and it will also serve as a benchmark for future mobility-related studies. Our study in this chapter reveals that the VMD outperforms MIPv6, HMIPv6, and PMIPv6 significantly, in terms of handoff latency, packet loss, and signaling overheadthe three important performance metrics to assess seamless user mobility as described in [39, 42, 110, 125, 126].

The rest of the chapter is organized as follows. In Section 4.1, the analytical models for the handoff-performance metrics are given for the VMD, MIPv6, HMIPv6, and PMIPv6 followed by the qualitative comparison of these protocols. Section 4.2 presents the assess-

[*] Portions of this chapter previously appeared as:
H. Tuncer, A. Kwasinski, and N. Shenoy, Performance Analysis of Virtual Mobility Domain Scheme vs. IPv6 Mobility Protocols, *Elsevier Computer Networks Journal*, Volume 57, Issue 13, 9 September 2013, Pages 2578-2596, ISSN 1389-1286.
H. Tuncer, Y. Nozaki, and N. Shenoy, Virtual domains for seamless user mobility, in *Proceedings of the 9th ACM international symposium on Mobility management and wireless access, ser. MobiWac11*. Miami, FL, USA, 2011.
H. Tuncer, Y. Nozaki, and N. Shenoy, Virtual Mobility Domains - a mobility architecture for the future internet, *Communications (ICC), 2012 IEEE International Conference on*, vol. 2774, no. 2779, pp. 10-15, June 2012.
H. Tuncer, Y. Nozaki, and N. Shenoy, Seamless user mobility in Virtual Mobility Domains for the future internet, *Communications (ICC), 2012 IEEE International Conference on*, vol. 5860, no. 5865, pp. 10-15, June 2012.

ment of the handoff-support performance of the VMD deployed in tier-1 cloud in an AS, while Section 4.3 focuses on intra-AS handoff support performance of the VMD deployed in tier-2 clouds in an AS. Section 4.4 analyzes the handoff performance of VMD deployed across multiple ASes and ISPs for inter-AS handoff support, followed by the summary of the performance study and the outcomes in Section 4.5.

4.1 Analytical Models

The handoff performance of VMD against MIPv6, HMIPv6, and PMIPv6 is analyzed using an analytical approach. The analytical study will be carried out for handoff latency and handoff signaling overhead. In Section 4.1.1, the different latency components that contribute to handoff latency are defined. Section 4.1.2 provides an overview of the components used in the signaling overhead calculation. The latency and signaling overhead equations used for MIPv6, HMIPv6, and PMIPv6 are presented in Section 4.1.3, 4.1.4, and 4.1.5, respectively.

4.1.1 Handoff Latency

Consistent with [39], we define *handoff latency* as the time that elapses between the events when a mobile node completes its layer 2 association at the new access router and receives a new data packet through the new access router. In general, the following delays are incurred during a handoff: movement detection delay, duplicate address detection delay, mobility control message transmission delay, AAA procedure delay, and first data packet transmission delay. They are defined as follows:

1. T_{MD} (Movement Detection delay) is the time spent for the discovery of a new access router performed through Router Solicitation/Advertisement (RS/RA) message exchanges between mobile node and access router [127]. Access Routers can be configured with minimum router advertisement interval (RAI_{min}) that is 0.5 s and maximum router advertisement interval (RAI_{max}) that is 1 s to send unsolicited router advertisement messages more often. Hence, the mean value of T_{MD} is half of the mean time between unsolicited router advertisement messages [39].

$$T_{MD} = (RAI_{min} + RAI_{max})/4 \tag{4.1}$$

   Detailed movement detection analysis can be found in [128].

2. T_{DAD} (Duplicate Address Detection delay) is the time spent by mobile node to ensure that a configured care-of address is likely to be unique on the new link. Duplicate address detection is performed by exchanging Neighbor Solicitation/Advertisement (NS/NA) messages. As per [36], T_{DAD} can be expressed as

$$T_{DAD} = RT \times DTimes \tag{4.2}$$

where *RT* is the time interval that mobile node waits after sending neighbor solicitation to see if a neighbor advertisement is forthcoming, and *DTimes* is the number of repetitions of solicit-and-wait process.

3. $t_{X,Y}$ is one-way transmission delay of a message between node X and node Y. The one-way transmission delay between mobile node and any router Y can be expressed as

$$t_{MN,Y} = (\frac{s}{B_{wl}} + L_{wl}) + (h_{MN,Y} - 1) \cdot (\frac{s}{B_w} + L_w + p_q) \tag{4.3}$$

where s is the message size (mobility related message sizes are provided in Table 4.1), $h_{MN,Y}$ is the number of hops between mobile node and Y, p_q is the average processing and queuing time of a packet at a router [129], B_{wl} and B_w are the bandwidth of wireless link and wired link respectively, L_{wl} and L_w are the propagation delay of wireless link and wired link respectively [125]. The round trip delay between X and Y is $2 \cdot t_{X,Y}$.

4. $T_{BU}(X, Y)$ denotes the time required to send a mobility control message to node X and to receive a response message in return. X can be mobile node in MIPv6 and HMIPv6; or mobility anchor gateway in PMIPv6. Y can be home agent or correspondent node in MIPv6, mobility anchor point as well in HMIPv6, or local mobility anchor in PMIPv6. $T_{BU}(X, Y)$ is calculated as follows

$$T_{BU}(X, Y) = 2 \cdot t_{X,Y} \tag{4.4}$$

5. T_{AAA} is the delay involved in mobile agents' communication with AAA servers to authenticate the mobile node.

6. T_{PT} is the data packet transmission delay from correspondent node to mobile node that includes the one-way transmission delay and the data packet interarrival time (t_P) in the worst case.

$$t_{CN,MN} \leq T_{PT} \leq t_{CN,MN} + t_P \tag{4.5}$$

Following assumptions have been made to conduct fair analysis across all mobility protocols.

1. The aforementioned four protocols are deployed in the same network topology under a single domain similar to the study in [39]. In MIPv6, the administrative domain under consideration is assumed to be a foreign network in which the mobile node is roaming. In HMIPv6, it is assumed to be the domain under a mobility anchor point. In PMIPv6, it is assumed to be domain under a local mobility anchor. In VMD, it is assumed to be a home network domain.

2. Handoff failure, link failure, and packet collision/back-off delay have not been considered [130]. We assume that every message is successfully transmitted to the destination, hence we are analyzing the normalized handoff delay as defined in [131].

3. We assume that the mobile nodes are allowed to access a network after AAA procedure is completed, and these access delays are the same for MIPv6, HMIPv6, PMIPv6, and VMD [39].

4. Layer 2 link switching delay is not included in the latency calculations and is in accordance with our latency definition.

The upcoming IPv6-based protocols' handoff latency equations are confirmed with [39].

4.1.2 Signaling Overhead

Signaling overhead incurred due to handoff is calculated as the sum of the mobility control messages multiplied by the number of hops that each message travels till they reach the destination node [42, 110]. The aim is to determine the number of bytes transmitted over the wired/wireless links due to handoff. $h_{X,Y}$ represents the number of hops between two nodes X and Y. In our calculations, we did not include router solicitation and router advertisement exchanges between mobile node and access router since they are not part of any mobility protocol. We also did not take into account the periodic binding refresh and the effect of a binding lifetime period as in [126]. Similarly, we excluded data packet delivery cost that takes into account data packet transmission, tunneling and data processing. The reason of excluding the data packet delivery is that data packets travel the same path in the network for all the protocols. All mobility control messages and their sizes are given in Table 4.1 [131].

The signaling overhead equations for IPv6-based protocols in the next sections are in confirmation with those in [42, 110].

4.1.3 Mobile IPv6

The equations for handoff latency and signaling overhead incurred in MIPv6 are presented in this section.

Table 4.1: The mobility control messages

Notation	*Description*	*Bytes*
BU_s	The binding update message	56
BA_s	The Binding Acknowledgement message s	56
BU_{cn_s}	The Binding Update message (to CN)	66
BA_{cn_s}	The Binding Acknowledgement message (from CN)	66
LBU_s	The Local Binding Update message	56
LBA_s	The Local Binding Acknowledgement message	56
PBU_s	The Proxy Binding Update message	76
PBA_s	The Proxy Binding Acknowledgement message	76
$HoTI_s$	The Home Test Init message	64
HoT_s	The Home Test message	74
$CoTI_s$	The Care-of Test Init message	64
CoT_s	The Care-of Test message	74
AAA_acc_s	The AAA acceptance message	326
AAA_req_s	The AAA request message	76
L_dereg_s	The location deregistration message	76

Handoff Latency in MIPv6

Handoff latency in MIPv6 (D_{MIPv6}) is expressed as follows using the terminology defined under Section 4.1.1 [42]:

$$\begin{aligned} D_{MIPv6} =& T_{MD} + 2 \cdot T_{DAD} + T_{AAA} + T_{BU}(MN, HA) + T_{RO} \\ &+ T_{BU}(MN, CN) + T_{PT} \end{aligned} \tag{4.6}$$

Duplicate address detection process happens first at the mobile node where the care-of address is created via stateless address autoconfiguration [36]. If the home agent does not have a binding for the care-of address, the home agent also performs the duplicate address detection on the care-of address before sending binding acknowledgement message to the mobile node. Stateless address autoconfiguration is one of the new features with IPv6 and does not require special communications (or protocols) with a DHCP server [132]. Hence, in this study it was decided to use the option of stateless address auto configuration, which then requires duplicate address detection to be executed in MIPv6 and HMIPv6.

T_{RO} represents the time spent for route optimization process which includes exchanging a sequence of signaling messages between mobile node, home agent, and correspondent node. Home address testing and care-of address testing can be initiated at the same time, hence T_{RO} is given by

$$T_{RO} = max(2 \cdot (t_{MN,HA} + t_{HA,CN}), 2 \cdot (t_{MN,CN})) \tag{4.7}$$

Signaling Overhead in MIPv6

Signaling overhead in MIPv6 (C_{MIPv6}) is expressed as follows [42]:

$$C_{MIPv6} = BU(MN, HA) + RO_{Home} + RO_{CN} + BU(MN, CN) \tag{4.8}$$

where the binding update/acknowledgement with home agent and correspondent node are represented by $BU(MN, HA)$ and $BU(MN, CN)$, respectively. They are given by

$$BU(MN, HA) = BU_s \cdot h_{MN,HA} + BA_s \cdot h_{HA,MN} \tag{4.9}$$

$$BU(MN, CN) = BU_{cn\_s} \cdot h_{MN,CN} + BA_{cn\_s} \cdot h_{CN,MN} \tag{4.10}$$

Route optimization process comprises two phases:

1. Exchanging a sequence of signaling messages between mobile node and home agent. The signaling overhead incurred during this phase is represented with RO_{Home} and it is calculated as

$$\begin{aligned} RO_{Home} = {} & HoTI_s \cdot h_{MN,HA} + HoTI_s \cdot h_{HA,CN} \\ & + HoT_s \cdot h_{CN,HA} + HoT_s \cdot h_{HA,MN} \end{aligned} \tag{4.11}$$

   where the messages are defined in Table 4.1.

2. Exchanging a sequence of signaling messages between mobile node and correspondent node through home agent. The signaling overhead incurred during this phase is represented by RO_{CN} and is given by

$$RO_{CN} = CoTI_s \cdot h_{MN,CN} + CoT_s \cdot h_{CN,MN} \tag{4.12}$$

   where the variables are defined in Table 4.1.

4.1.4 Hierarchical Mobile IPv6

The handoff latency and signaling overhead incurred in HMIPv6 are formulated in this section.

Handoff Latency in HMIPv6

If a mobile node moves to an HMIPv6 domain for the first time, initial HMIPv6 latency (D_{HMIPv6_o}) is expressed as follows [39]:

$$\begin{aligned} D_{HMIPv6\_o} = {} & T_{MD} + 3 \cdot T_{DAD} + T_{AAA} + T_{BU}(MN, MAP) + \\ & T_{BU}(MN, HA) + T_{RO} + T_{BU}(MN, CN) + T_{PT} \end{aligned} \tag{4.13}$$

T_{DAD} is multiplied by three to include duplicate address detection delay for on-link care-of address at the mobile node and the regional care-of address at the mobility anchor point and at the home agent. However, while the mobile node moves within the HMIPv6 domain [39], the latency D_{HMIPv6} is given by

$$D_{HMIPv6} = T_{MD} + T_{DAD} + T_{AAA} + T_{BU}(MN, MAP) + T_{PT} \tag{4.14}$$

where T_{DAD} represents the duplicate address detection delay for on-link care-of address.

Signaling Overhead in HMIPv6

Signaling overhead in HMIPv6 (C_{HMIPv6}) caused by local binding update/acknowledgement messaging between the mobile node and the mobility anchor point [110] is formulated as

$$C_{HMIPv6} = BU(MN, MAP) \tag{4.15}$$

where $BU(MN, MAP)$ represents the signaling overhead incurred due to binding update messaging between mobile node and mobility anchor point and is calculated as

$$BU(MN, MAP) = LBU_s \cdot h_{MN,MAP} + LBA_s \cdot h_{MAP,MN} \tag{4.16}$$

4.1.5 Proxy Mobile IPv6

In this section, handoff latency and signaling overhead equations for PMIPv6 are presented.

Handoff Latency in PMIPv6

If a mobile node moves to a PMIPv6 domain for the first time, MIPv6 will be used to support macro mobility. The handoff latency (D_{PMIPv6_o}) is thus represented as follows [39]:

$$\begin{aligned} D_{PMIPv6\_o} =& T_{AAA} + T_{BU}(MAG, LMA) + T_{DAD} + \\ & T_{BU}(MN, HA) + T_{RO} + T_{BU}(MN, CN) + T_{PT} \end{aligned} \tag{4.17}$$

As seen in (4.17), movement detection is not required within a PMIPv6 domain because of the mobility access gateway [133].

Mobile node's handoff latency within the PMIPv6 domain (D_{PMIPv6}) [39] is calculated as follows:

$$D_{PMIPv6} = T_{AAA} + T_{BU}(MAG, LMA) + T_{PT} \tag{4.18}$$

Duplicate address detection does not occur because the address assigned by local mobility anchor is used by mobile node in the entire PMIPv6 domain [38].

Signaling Overhead in PMIPv6

The signaling overhead in PMIPv6 (C_{PMIPv6}) is incurred because of the communication of local mobility anchor with the new mobility access gateway (MAG_n) and the old mobility access gateway (MAG_o) for registration or deregistration of the mobile node, respectively [110].

$$C_{PMIPv6} = BU_{reg}(MAG_n, LMA) + BU_{dereg}(MAG_o, LMA) \tag{4.19}$$

where $BU_{reg}(MAG_n, LMA)$ represents signaling overhead that is incurred during registration of mobile node at MAG_n and is given by

$$BU_{reg}(MAG_n, LMA) = PBU_s \cdot h_{MAG_n,LMA} + PBA_s \cdot h_{LMA,MAG_n} \tag{4.20}$$

$BU_{dereg}(MAG_o, LMA)$ represents the signaling overhead that is incurred during deregistration of mobile node at the MAG_o and it is formulated as follows:

$$BU_{dereg}(MAG_o, LMA) = PBU_s \cdot h_{MAG_o,LMA} + PBA_s \cdot h_{LMA,MAG_o} \tag{4.21}$$

4.1.6 Virtual Mobility Domain

In this section, we provide the equations for handoff latency and signaling overhead that are incurred in VMD.

Handoff Latency in VMD

Intra-domain handoff delay in VMD (D_{VMD}) includes AAA procedure delay and binding update messaging between access router and common anchor cloud of the old access router and the new access router. In VMD, movement detection is not required as it is a network-based mobility management scheme. Duplicate address detection is also not required because VMD assigns a unique address to a mobile node as explained in Section 3.3.1. Hence, D_{VMD} is given by

$$D_{VMD} = T_{AAA} + T_{BU}(AR, CAC) + T_{PT} \tag{4.22}$$

where CAC is the *common anchor cloud* between the old access router and new access router. For example, if mobile node enters the VMD for the first time, or mobile node moves between the tier-2 clouds in Fig. 3.2, the CAC will be Root Cloud. If the mobile node moves between access routers connected to the same tier-2 cloud, then the CAC will be that tier-2 cloud.

Signaling Overhead in VMD

During a mobile node's roaming in VMD, wired nodes (i.e., access routers, and routers in the tier-1 and tier-2 network clouds) handle the mobility management. The signaling overhead is incurred because of the location update messaging between access router and the different entities in the VMD cloud. Signaling overhead in VMD (C_{VMD}) is expressed as follows:

$$C_{VMD} = BU_{reg}(AR_n, CAC) + BU_{dereg}(AR_o, CAC) \tag{4.23}$$

where CAC is the common anchor cloud between the previous access router (AR_o) and the new access router (AR_n).

$BU_{reg}(AR_n, CAC)$ represents the signaling overhead that is incurred during registration of the mobile node to AR_n which requires AAA_req and AAA_acc message exchanges between AR_n and CAC. $BU_{reg}(AR_n, CAC)$ is expressed as

$$BU_{reg}(AR_n, CAC) = AAA\_req_s \cdot h_{AR_n,CAC} + AAA\_acc_s \cdot h_{CAC,AR_n} \tag{4.24}$$

$BU_{dereg}(AR_o, CAC)$ represents the signaling overhead that is incurred to deregister mobile node at AR_o which requires L_dereg message transmission between AR_o and CAC. $BU_{dereg}(AR_o, CAC)$ is calculated as follows

$$BU_{dereg}(AR_o, CAC) = L\_dereg_s \cdot h_{CAC,AR_o} \tag{4.25}$$

4.1.7 Operational Comparison

In this section, we discuss the operational differences between VMD and the IPv6-based mobility protocols. Table 4.2 presents the features of these protocols for easy comparison.

VMD vs. MIPv6

VMD is different from IPv6-based mobility protocols as it is based on a new FCT internetworking model that does not use IP addressing and IP-based routing protocols. In the mobility management scheme, a mobile node uses only one address assigned under a VMD cloud. After a mobile node is registered under the VMD, data packet forwarding between a correspondent node and a mobile node is handled by the FCT packet forwarding process as explained in Section 3.3.4 and does not require route optimization. A correspondent node directly sends data packets to the VMD. The network clouds in the VMD forward data packets to the mobile node in accordance with their forwarding bases. During the movement of the mobile node in the same VMD, only the mobile node entry in the related forwarding bases is updated with a new downland address as stated

Table 4.2: Qualitative comparison of MIPv6, HMIPv6, PMIPv6, and VMD

Category	*MIPv6*	*HMIPv6*	*PMIPv6*	*VMD*
Mobility Management	Host-based	Host-based	Network-based	Network-based
Network Architecture	Flat	Hierarchical	Hierarchical	Tiered
Target Network	IP	IP	IP	FCT
Operating OSI Layer	3	3	2 & 3	2 & 3
Required Infrastructure	HA	AR & MAP	LMA & MAG	VMD
Router Advertisement	Broadcast	Broadcast	Unicast	Unicast
Addressing Model	Shared-prefix	Shared-prefix	Per-MN-prefix	Shared-prefix
MN Address	HoA	RCoA & LCoA	CoA	Tiered address
Address Type	IPv6	IPv6	IPv6	Tiered
Address Length (bits)	128	128	128	Dynamic
Address Change in Domain	Yes	Yes	No	No
Address Assigned by	HA	MAP	LMA	VMD
Tunneling	Inter-AS	Intra-AS	Intra-AS	No
Duplicate Address Detection	Yes	Yes	No	No
Route Optimization	Yes	Yes	Yes	Yes

in Section 3.3. Therefore, updating a correspondent node is not necessary. VMD supports a network-based mobility management scheme and wired nodes handle the mobility management on behalf of the mobile node as also expressed in (4.22) and (4.23). This results in reduced latency, signaling overhead, and wireless resource usage as described in Chapter 4.

VMD vs. HMIPv6

In HMIPv6, the node movement within the HMIPv6 domain is not visible outside of the domain. However, HMIPv6 accomplishes this by using two addresses for a mobile node. As expressed in (4.14), HMIPv6 requires duplicate address detection and the mobile node is expected to involve in mobility management as HMIPv6 supports a host-based mobility management while VMD supports a network-based mobility management. HMIPv6 uses tunneled data communications which requires higher signaling overhead during data packet transmission for a mobile node as compared to VMD.

VMD vs. PMIPv6

VMD and PMIPv6 are both network-based mobility management schemes. These protocols thus avoid movement detection and duplicate address detection processes as seen in

(4.18) and (4.22). Different from PMIPv6, VMD provides the collaborative mobility management which limits the effect of the mobility related signaling to the network clouds that have mobile node profile in their proxy AAA servers as explained in Section 3.2. This benefits in fewer message exchanges and lower latency as discussed in the next section. PMIPv6 still requires tunneling for mobile node's data communication.

4.2 Tier-1 Deployment of the Protocols in an AS for Intra-AS Roaming

In this section, we show the analytical results based on the equations derived in Section 4.1. The mobility protocols are deployed in the network illustrated in Fig. 4.1, which is the same network of Fig. 3.2. We implemented VMD, MIPv6, HMIPv6 and PMIPv6 on OPNET based on related RFC documents [35,38] to validate the analytical results. In the case of HMIPv6, mobility anchor point is located at Root Cloud. For PMIPv6, local mobility anchor is located at Root Cloud while the mobility access gateway is located at each access router. To conduct comparison with MIPv6, the network in Fig. 4.1 will be assumed to be a foreign network, and hence home network and home agent are located outside of the AS.

The AS network topology that is used for analytical and simulation studies is shown in Fig. 4.1. Network settings are based mainly on those provided in [97] and OPNET modeler default settings and decided primarily based on RFCs. Access Routers operate using 802.11g with a data rate of 54 Mbps. Access Routers send layer-2 beacons every 20 milliseconds (ms) while router advertisements are uniformly distributed between 0.5 seconds (s) and 1 s as per OPNET modeler default settings. The transmit power is 0.05 watts with a power threshold of -80 dBm. "WLAN Beacon Efficiency Mode" in OPNET is disabled to have real beacon transmission on wireless media. Coverage areas of neighboring access routers overlap to avoid link breaks at the physical layer. Wired links have a data rate of 5 Mbps, similar to [97]. To include the effect of distances, link L1 has propagation delay of 2 ms, link L2 4 ms, and link L3 10 ms, similar to [97]. Wireless link propagation delay is assumed negligible as it is in magnitudes of nanoseconds in this network setup. We used the same propagation delay values in calculating (4.3) for the analytical results. The data traffic from the correspondent node to the mobile node has uniform packet interarrival time of 10 ms, because each router processes a data packet every 10 ms.

The mobile node travels along the trajectory shown by the red line (in Fig. 4.1) at a speed of 30 km/h. This path is chosen because along the path the mobile node performs two different handoff types that we aim to observe, which are intra-cloud handoff (i.e. handoff 1 and handoff 3) and inter-cloud handoff (i.e. handoff 2 and handoff 4). At each

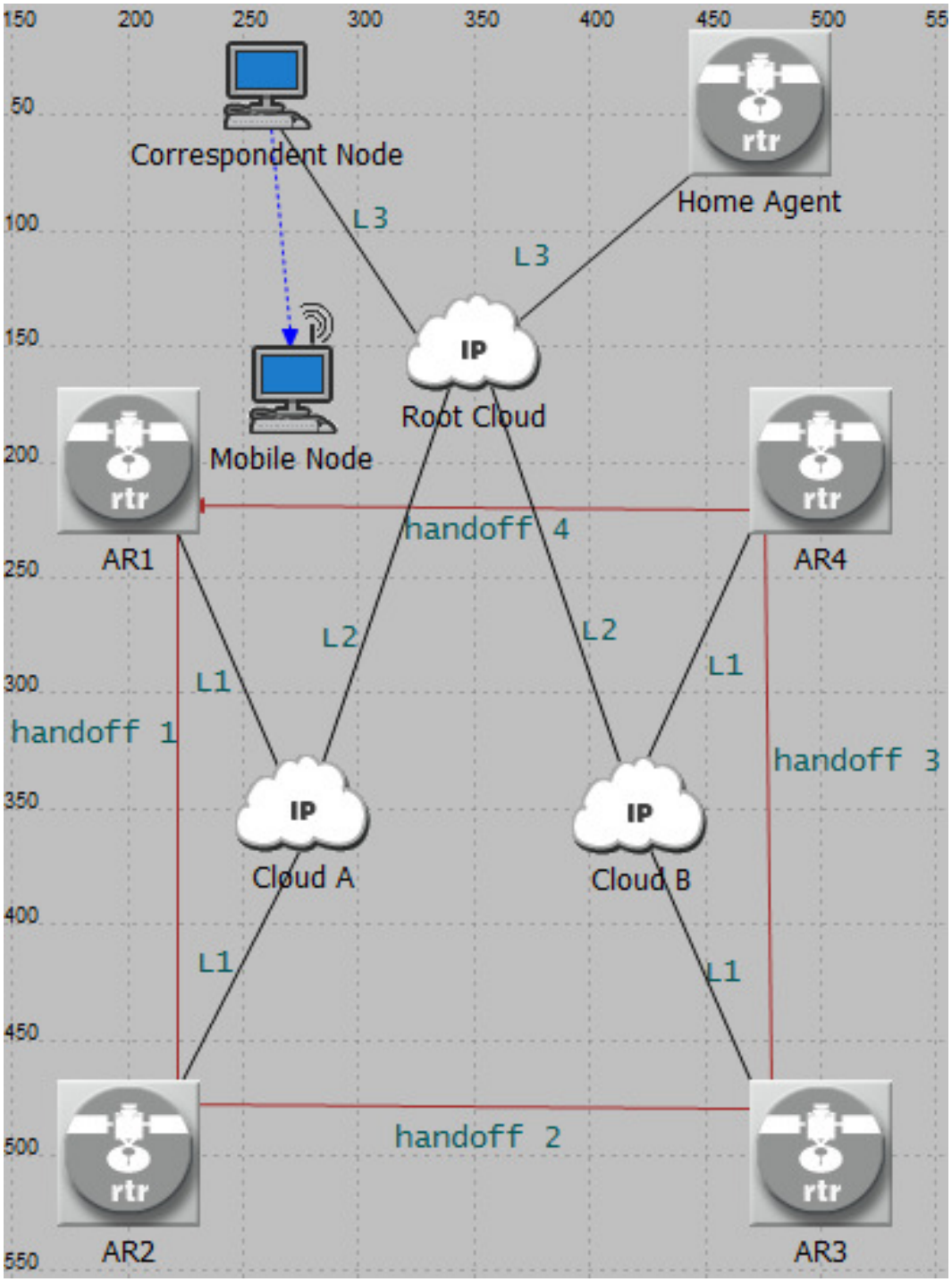


Figure 4.1: The network topology that is used for OPNET simulation and analytical study.

access router, the mobile node waits for 2 minutes before moving to the coverage area of the next access router. The recorded readings were averaged over 20 simulation runs with different seeds and the standard deviation of the latency results are 0.3 s which does not change significantly when the number of simulation runs are increased. As the focus is on handoff performance, the investigations and results are limited to a single node that goes through the handoff process.

It was verified that the MIPv6, HMIPv6, and PMIPv6 simulation performance complied with that described in [97, 134]. Simulation logs for each protocol were carefully examined to see that each protocol executes all the processes stated in Section 4.1

The performance of the aforementioned protocols are analyzed mainly according to handoff latency, signaling overhead, and packet loss metrics. The detailed performance analysis is given in the next sections.

4.2.1 Handoff Latency

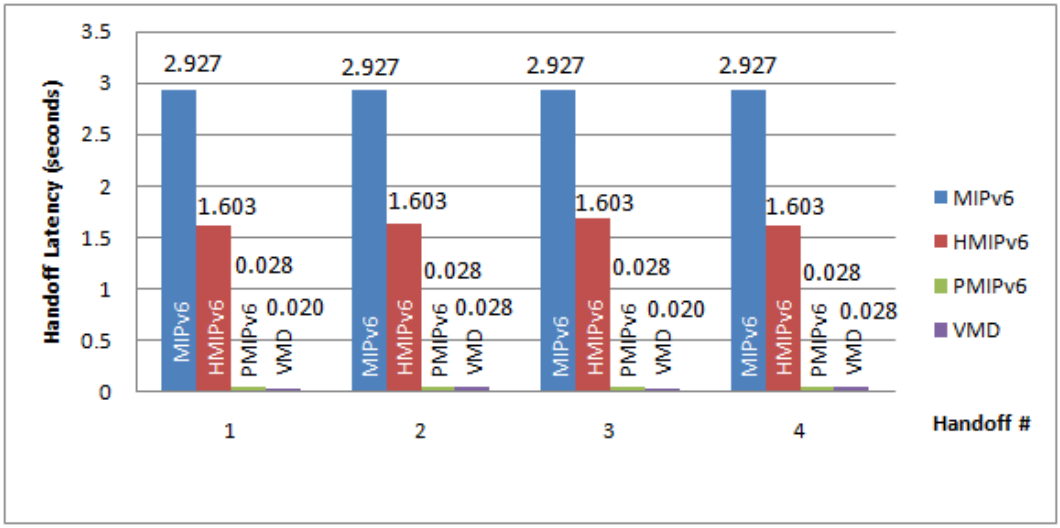


Figure 4.2: Analytical handoff latency results for MIPv6, HMIPv6, PMIPv6 and VMD.

Fig. 4.2 shows the latency results based on the equations presented in Section 4.1. Fig. 4.3 presents the latency values recorded using OPNET simulations. We observed a lower than 5% difference between analytical and simulation latency results in most cases. The higher latency results in the simulation are caused by the varying values of movement detection T_{MD} and duplicate address detection T_{DAD} mainly because these processes are

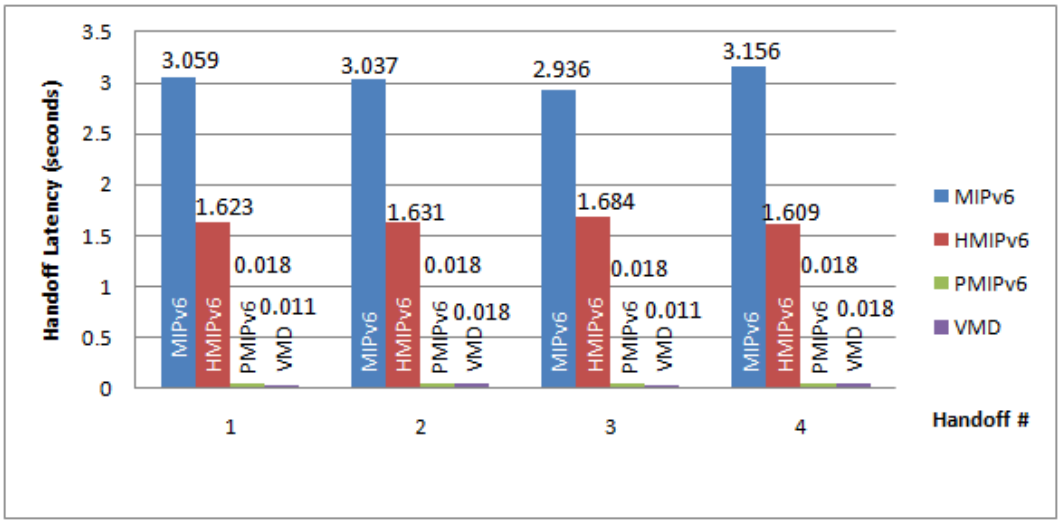


Figure 4.3: Handoff latency results for MIPv6, HMIPv6, PMIPv6 and VMD observed in OPNET simulations.

modeled as random processes by OPNET while the analytical results are based on the analytical models of Section 4.1 and representative of those discussed in [39, 110, 125].

In OPNET simulation, router advertisements and movement detection are handled by *ipv6_ra_host* module. The mobile node considers a new router advertisement, that comes from the new access router as a movement hint. In OPNET simulation logs, we observed that T_{MD} gets values between 0.2 and 0.8 s while T_{MD} value used in analytical results is derived from (4.1).

In analytical papers, varying values are used for duplicate address detection delay (T_{DAD}): 0.5 s [126] and 1 - 2 s [135]. In OPNET default MIPv6 network and our implementation, we observed that T_{DAD} gets varying values between 1 s and 1.4 s. In our analytical study, we assign T_{DAD} 1.2 s in accordance with (4.2).

In the simulation, we also observed that the mobile node receives data packet from correspondent node around 6 - 8 ms after the mobility related messaging is completed. This is less than T_{PT} value expressed in (4.5). From the simulation logs, we observed that mobility management is completed just before data packet arrives to the AS network, which means that while the data packet travels between correspondent node and Root Cloud (which takes 10 ms), mobility management is in process. All the above differences in the OPNET modeler and the equations result in a difference of 9-10 ms which is approximately 5%.

The comparative performance of all the four protocols are however similar in the analytical and simulation studies. MIPv6 has the highest handoff latency across all the four handoff events - typically 2.927 s from the analytical models and around 3 s from the simulation models. This is because MIPv6 is a macro-mobility protocol and has no optimized operations for roaming across subnets within a mobility domain. The mobile node has to inform its home network and execute route optimization-related signaling with correspondent node every time it changes its access router. Furthermore, movement detection and duplicate address detection procedures have to be executed for each mobile node movement as expressed in (4.6). MIPv6 latency plots however are useful as benchmark against which the performance of the other three protocols can be compared.

HMIPv6 has reduced handoff latency, around 1.6 s as the global identifier of the mobile node, which is its regional care-of address, never changes in the HMIPv6 domain. Thus, the mobile node does not have to inform either the home network or the correspondent node when the mobile node connects to another access router. However, the mobile node has to create its on-link care-of address and execute duplicate address detection for the new on-link care-of address when it connects to a new access router inside the mobility anchor point domain. After it successfully decides on the on-link care-of address, it has to register its on-link care-of address to the mobility anchor point. As expressed in (4.14), the mobile node's involvement in the on-link care-of address creation, binding update to the mobility anchor point, and the wireless media usage in these processes explain the higher latency of HMIPv6 compared to PMIPv6 and VMD.

PMIPv6 and VMD have very low latencies, in the order of milliseconds compared to MIPv6 and HMIPv6 mainly because PMIPv6 and VMD are network-based mobility management protocols. Therefore, movement detection and duplicate address detection procedures are avoided in (4.18) and (4.22). They also do not require mobile node's involvement in mobility management and this avoids communication over wireless links. Mobility management is handled by the wired nodes. PMIPv6 and VMD have the same mobility performance for inter-cloud movement (handoff 2 and handoff 4) of the mobile node since mobility messaging occurs between access router and Root Cloud in VMD, between mobility access gateway and local mobility anchor in PMIPv6 which incur same delay. However, during intra-cloud roaming (handoff 1 and handoff 3) VMD handoff latency is lower than PMIPv6. This is because in VMD the mobile node movement is handled by the mobility agent in the closest common anchor cloud between the old and new access routers that the mobile node is moving across. For handoff 1 and handoff 3, the common anchor cloud is tier-2 cloud. Due to collaborative mobility management, VMD performs better than PMIPv6 for the cases where the common anchor cloud is lower than the cloud under which the VMD was deployed. However, in the case of PMIPv6, a

mobility access gateway has to communicate with a local mobility anchor for all handoffs.

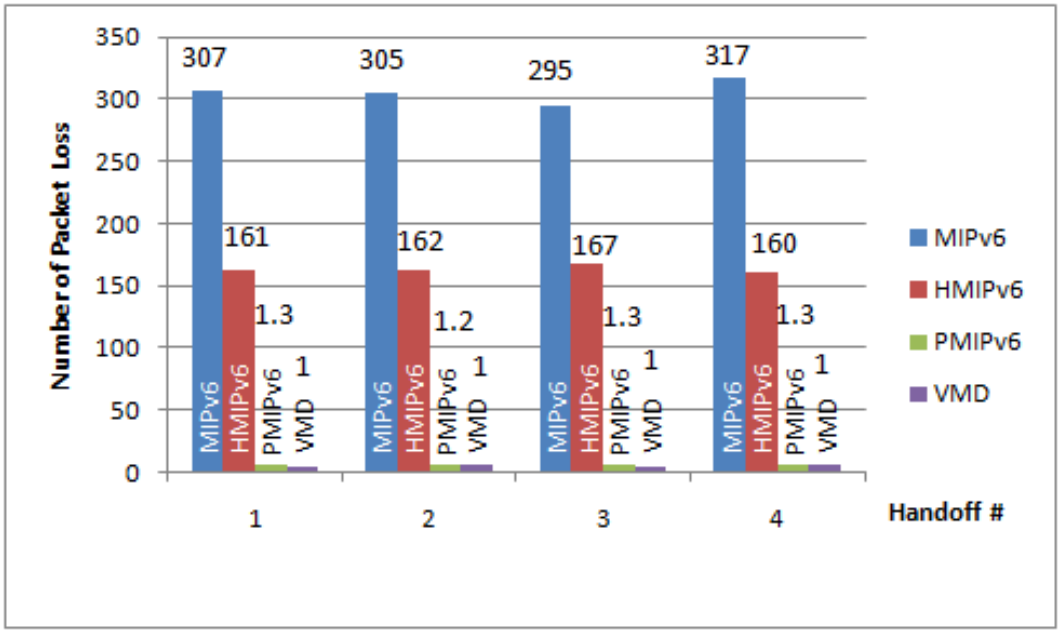


Figure 4.4: Number of packets lost during the four handoff events observed in OPNET simulations.

4.2.2 Packet Loss

Fig. 4.4 shows the packet loss during the four handoff events in OPNET simulation. The results show a trend similar to that observed with handoff latencies in Fig. 4.3. With a constant interarrival time between packets that are sent from the correspondent node to the mobile node, the packet loss will be proportional to the handoff latency. MIPv6 and HMIPv6 cause more than 290 and 160 packet loss, respectively. In PMIPv6 and VMD, the packet loss is around 1.

4.2.3 Signaling Overhead

Fig. 4.5 is the plot of signaling overhead in bytes under the four handoff events and the four mobility protocols from analytical models. Fig. 4.6 provides similar plots from the OPNET simulation. The analytical and simulation results are identical. MIPv6 has a very high signaling overhead of 2356 bytes, because binding update and acknowledgements are to be exchanged between the mobile node and its home agent - in this case the home agent is considered outside the mobility domain. As given in (4.11) and (4.12), the route

optimization process also requires several message exchanges between the mobile node, the home agent, and the correspondent node.

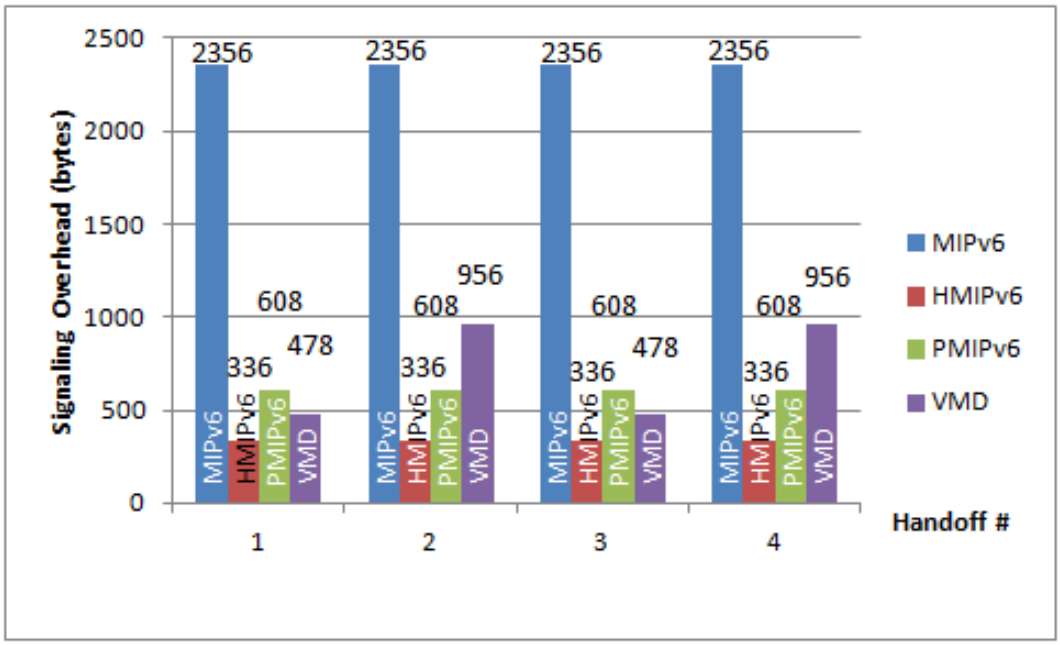


Figure 4.5: Analytical signaling overhead results during handoffs.

The signaling overhead with HMIPv6, PMIPv6, and VMD are less than 1000 bytes because the AS network has been defined as the mobility domain and any signaling is thus constrained to this mobility domain. In the case of HMIPv6, the signaling overhead is due to the exchange of binding update and acknowledgement messages between the mobile node and the mobility anchor point as in (4.15). No signaling is required between the mobile node and the home agent, and the mobile node and the correspondent node as the global address of mobile node remains unchanged. PMIPv6 has 608 bytes overhead because the binding update and acknowledgement messages are longer as seen in Table 4.1 and a deregistration message from the mobility access gateway of the previous access router to the local mobility anchor is required as expressed in (4.19). In the case of VMD, the signaling overhead is 476 bytes during handoff 1 and handoff 3 as the handoff is managed locally within the tier-2 cloud, where the cloud routers communicate with access routers. However, for handoff 2 and handoff 4 the signaling overhead is 956 bytes as communication among the tier-2 clouds and Root Cloud is required for the handoff. The reason of having higher signaling overhead is due to the size of AAA acceptance message. AAA acceptance message is 326 bytes long since it carries mobile node profile

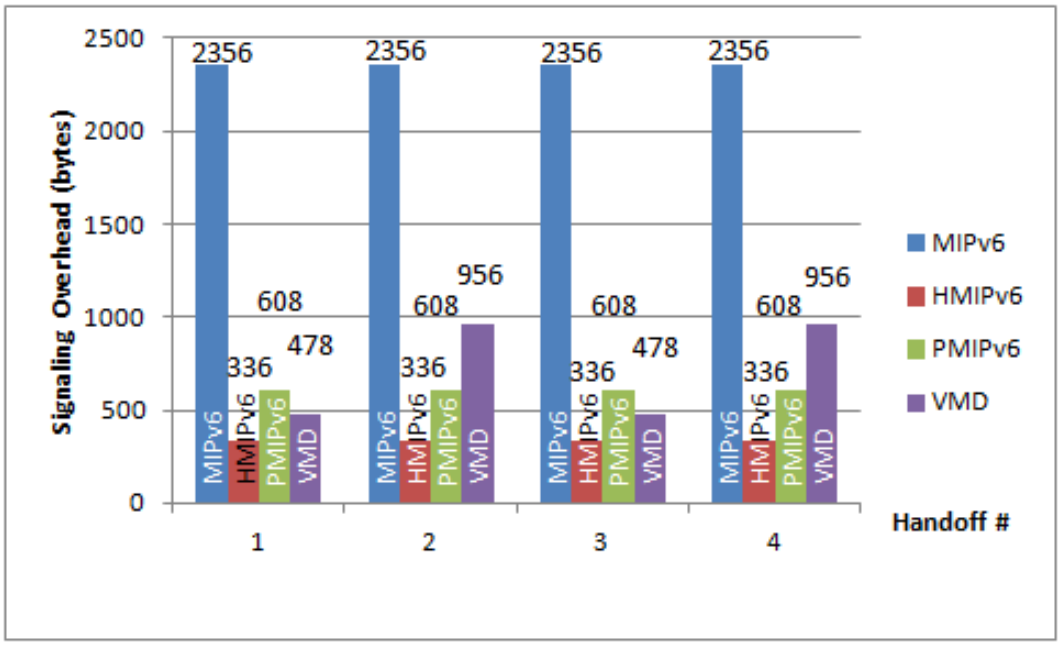


Figure 4.6: Signaling overhead during handoffs observed in OPNET simulations.

which is 250 bytes long [136].

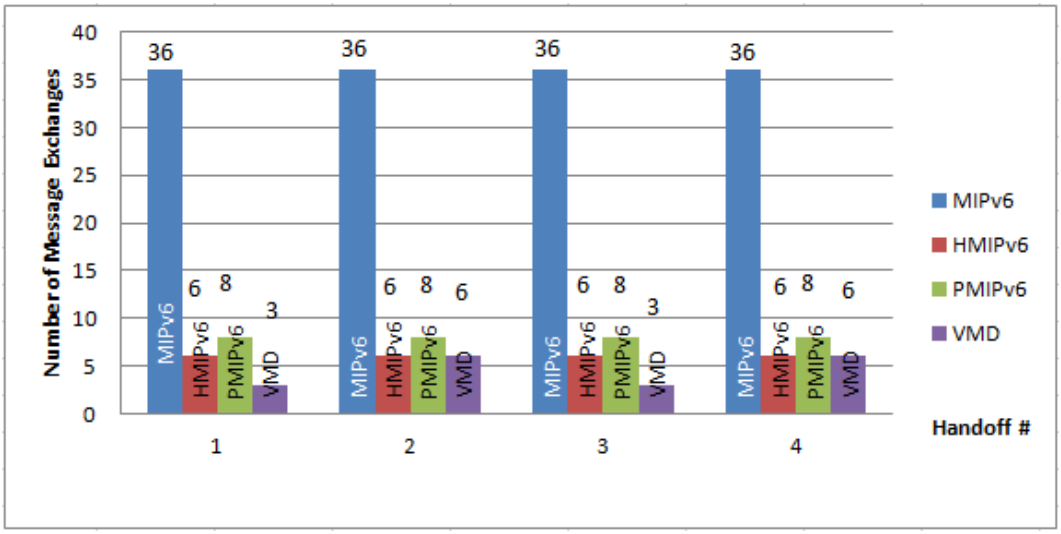


Figure 4.7: Analytical number of message exchange results during handoffs.

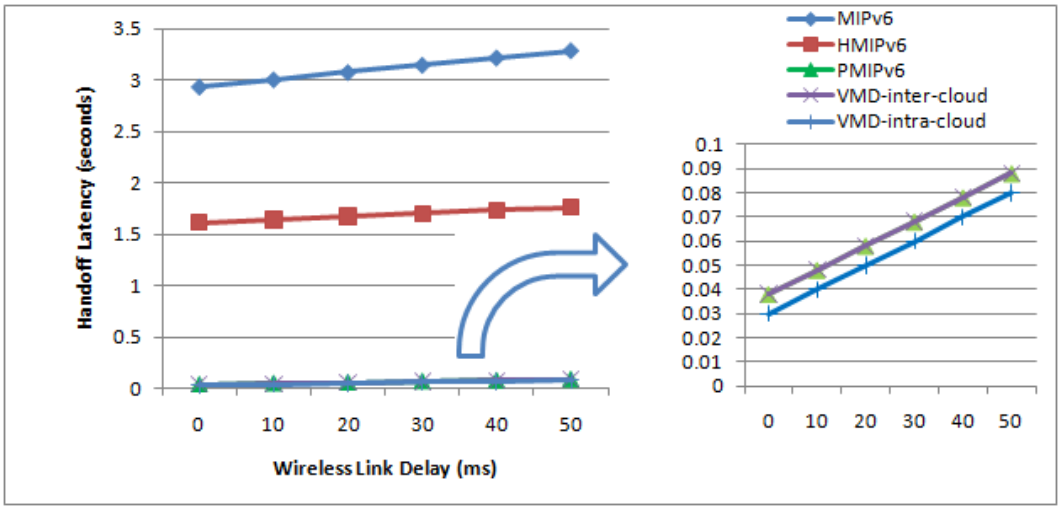


Figure 4.8: Analytical results for impact of wireless link delay over handoff latencies in MIPv6, HMIPv6, PMIPv6, and VMD (intra-cloud & inter-cloud).

Number of mobility control message exchanges among the nodes are important because each message exchange requires processing which may affect the handoff performance. Fig. 4.7 shows that MIPv6 has the highest number of message exchanges due to the binding update and route optimization between home agent and correspondent node as expressed in (4.8). HMIPv6 causes 6 message exchanges due to binding update process while PMIPv6 causes 8 message exchanges due to additional deregistration process as expressed in (4.15) and (4.19), respectively. VMD executes 3 message exchanges for intra-cloud handoff and 6 message exchanges for inter-cloud handoff as expressed in (4.23). In VMD, a mobile node does not involve in message exchanges, only access routers and the common anchor cloud which causes fewer message exchanges.

4.2.4 Factors Affecting Handoff Latency

We provided overall comparison of handoff latency performances of the protocols in Section 4.2.1. Mobility protocols' performance is sensitive to wireless/wired link quality, network density, and network setup which may vary in real-world situations. Therefore, we investigate how mobility protocols get affected from changing wireless/wired link delays and movement detection delays. Our analysis is based on the analytical models in Section 4.1, as it is convenient to adjust the parameters contributing to the handoff latency.

Wireless Link Effect on Handoff Latency

Wireless communication delay comprises transmission delay, propagation delay, and access latency depending on network density, wireless medium, and communication technology. In Fig. 4.8, we present the effect of varying wireless communication delay, namely wireless link delay, on the mobility protocols' handoff performance. The network setup and all parameters explained at the beginning of Section 4.2 are maintained. As seen in Fig. 4.8, handoff latencies in each protocol increases with different magnitudes. MIPv6 is the most affected protocol with 70 ms increase for the 10 ms increment step in the wireless link delay. As expressed in (4.6), the mobile node involves in binding update with the home agent and the correspondent node ($T_{BU}(MN, HA)$ and $T_{BU}(MN, CN)$ respectively), and also route optimization (T_{RO}) procedures which require a high level of message exchange over wireless medium. HMIPv6 latency increases by 30 ms for each 10 ms increment on wireless link delay since the mobile node only does binding update with the mobility anchor point once. On the other hand, PMIPv6 and VMD are affected from wireless link delay changes only because of the data packet transmission (T_{PT}), which is also present in all protocols' latency calculation. The bottom right of Fig. 4.8 shows the enlarged view of the PMIPv6 and VMD results. PMIPv6 and VMD inter-cloud handoff latencies continue to overlap as explained in the previous section. The difference between VMD intra-cloud handoff and inter-cloud handoff stays the same, which is 8 ms caused by the round-trip over the wired link between the tier-2 cloud and the Root Cloud.

Wired Link Effect on Handoff Latency

Wired communication may be affected by various factors such as the network congestion, quality of the wired links, and distance between nodes. The varying performance of wired communication is represented with wired link delay in Fig. 4.9 and 4.10, where we aim to capture the effect of wired communication delay on mobility protocols' performance. Our analysis is twofold: (i) the links in the mobility domain and (ii) the links outside the mobility domain, mainly to observe performance of micro-mobility and macro-mobility protocols distinctly.

The wired link delay between access router and Root Cloud namely, L1 and L2 in Fig. 4.1 are set identical and are given varying values to observe the effect of wired link delay in the mobility domain. All the other parameters stated at the beginning of Section 4.2 are kept the same. Fig. 4.9 depicts that MIPv6 is the most affected protocol with 70 ms increase in handoff latency for each 10 ms increment on wired link delay, since binding messages and route optimization messages are transferred through the wired links between access router and Root Cloud. On the other hand, handoff in HMIPv6, PMIPv6, and VMD-intra-cloud increase by 30 ms for a 10 ms step-up on wired link delay since only the binding messages are transferred on the link between access router and Root

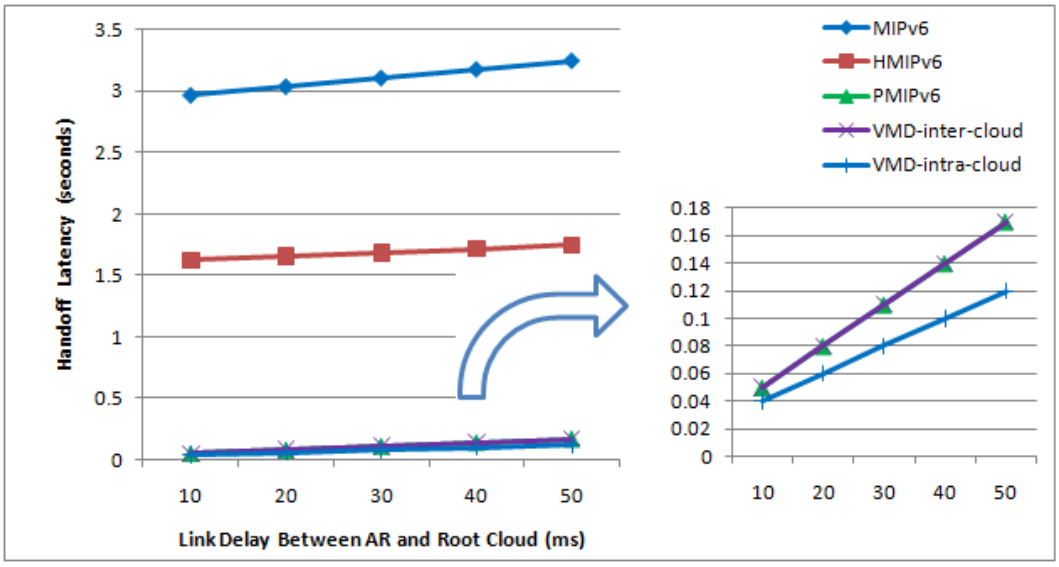


Figure 4.9: Analytical results for impact of wired link delay between access router and Root Cloud over handoff latencies in MIPv6, HMIPv6, PMIPv6, and VMD (intra-cloud & inter-cloud).

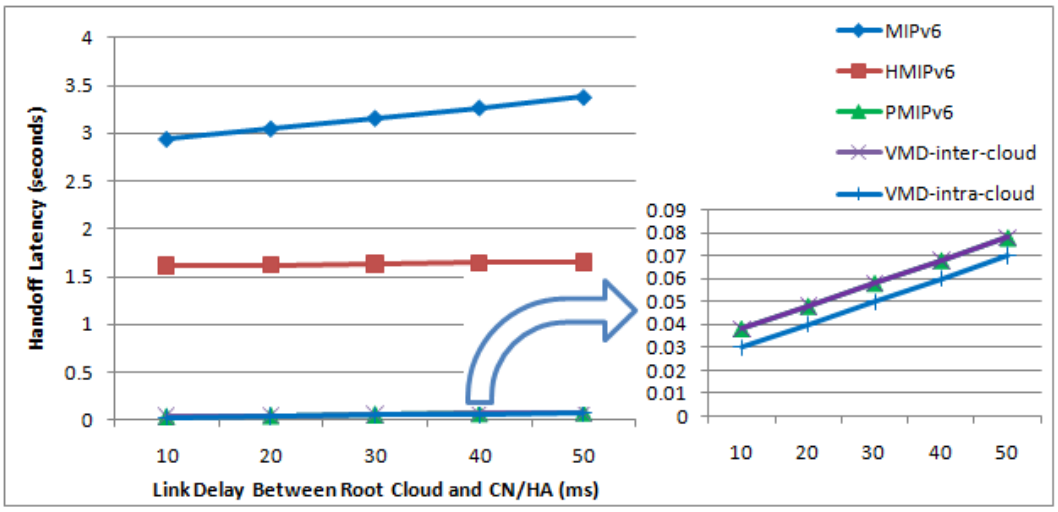


Figure 4.10: Analytical results for impact of wired link delay between Root Cloud and correspondent node or home agent over handoff latencies in MIPv6, HMIPv6, PMIPv6, and VMD (intra-cloud & inter-cloud).

Cloud. However, intra-cloud handoffs in the VMD gets affected least (20 ms), since the

mobility related messaging only occur between access router and tier-2 cloud, and not Root Cloud.

We next vary delays on the links between Root Cloud and home agent/correspondent node, namely L3 in Fig. 4.1 to show the effect of communicating with correspondent node or home agent that are outside of the mobility domain. As in Fig. 4.10, MIPv6 is most affected with 110 ms increase in handoff latency for 10 ms increment on the link delay since MIPv6 is a macro-mobility management protocol and the mobile node informs home agent and correspondent node for every movement. However, micro-mobility protocols do not require informing home agent or correspondent node because the intra-domain movements of the mobile node are only visible to local domain and handled by mobility anchor point, local mobility anchor or VMD depending on the protocol. In PMIPv6, HMIPv6, and VMD, the increase in handoff latency is only due to the data packet transmission from correspondent node.

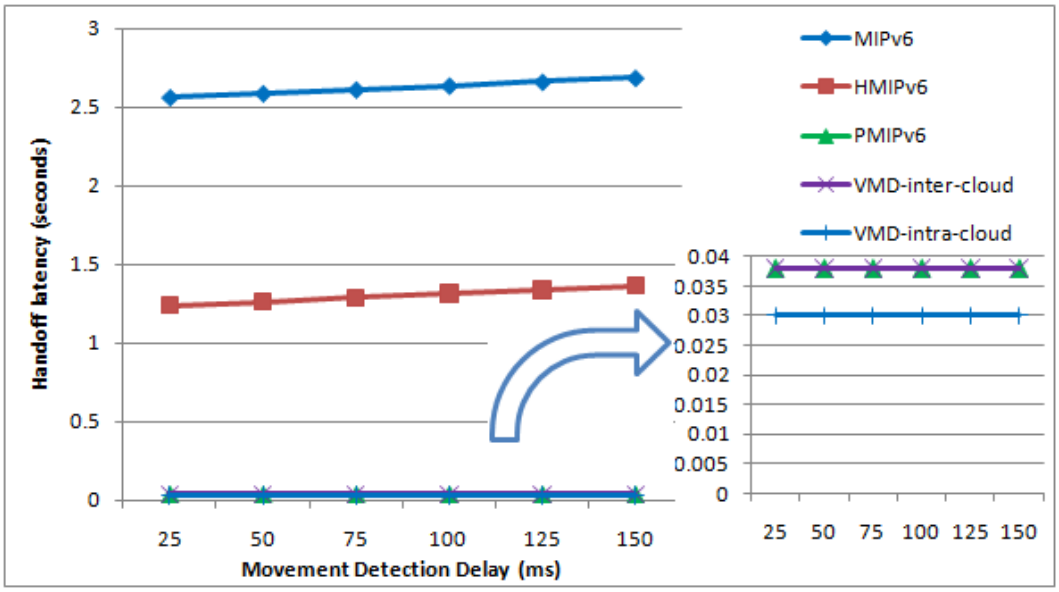


Figure 4.11: Analytical results for effect of movement detection delay over the handoffs in mobile IP protocols and VMD.

Movement Detection Effect on Handoff Latency

Movement detection delay may depend on several factors such as access router configuration, router advertisement interval time, wireless medium, and access technology. Fig.

4.11 presents the impact of varying T_{MD} over the handoff latencies. In PMIPv6 and VMD, layer 3 movement detection is not required since these protocols deploy mechanisms to control mobile node's attachment or detachment such as layer-2 triggers. Therefore, from the perspective of the mobile node, the VMD or the PMIPv6 domain appears as a home network. On the other hand, MIPv6 and HMIPv6 handoff delays increase by the increase in T_{MD}. MIPv6 and HMIPv6 require movement detection of a mobile node as they are host-based mobility protocols and the mobile node has to get new address at the new access router.

4.3 Tier-2 Deployment of the Protocols in an AS for Intra-AS Roaming

This section presents the simulation results for the extended deployment of VMD and IPv6-based protocols to tier 2 in the same AS network presented in Section 4.4. Mobility anchor point (in HMIPv6), local mobility anchor (in PMIPv6) and virtual cloud (in VMD) are deployed in tier-2 clouds i.e. Cloud A and B on the network illustrated in Fig. 4.1. In the case of MIPv6, home network and home agent are located outside of the AS the same as Section 4.4. The other network settings are kept the same with Section 4.4. The aim is to analyze the impact of mobility domain size on aforementioned protocols' performance in OPNET simulations. The collected results are compared to the scenario where the protocols are deployed to tier-1 clouds, stated in Section 4.4.

4.3.1 Handoff Latency

In Fig. 4.12, MIPv6 has the same latency values with the Fig. 4.3 that is observed in the scenario in Section 4.4 as MIPv6 deployment is not changed. HMIPv6 has latency of 1.6 s. This is only few ms less than the tier-1 deployment of the HMIPv6 in Fig. 4.3. In the tier-2 deployment scenario, binding messages do not travel the link between tier-2 and tier-1 clouds (has 4 ms delay) because mobility anchor point is deployed in a tier-2 cloud, rather than the tier-1 cloud. However, movement detection and duplicate address detection processes are dominant contributors to the latency with magnitudes of seconds. PMIPv6 has reduced latency that is 11 ms because local mobility anchor is deployed in a tier-2 cloud and binding messages are sent through access routers and tier-2 clouds. On the other hand, the VMD performs the same as in the tier-1 deployment scenario regardless of the domain size because mobility management is handled by the tier-2 cloud in both scenarios with the help of the collaborative mobility management scheme introduced via VMD.

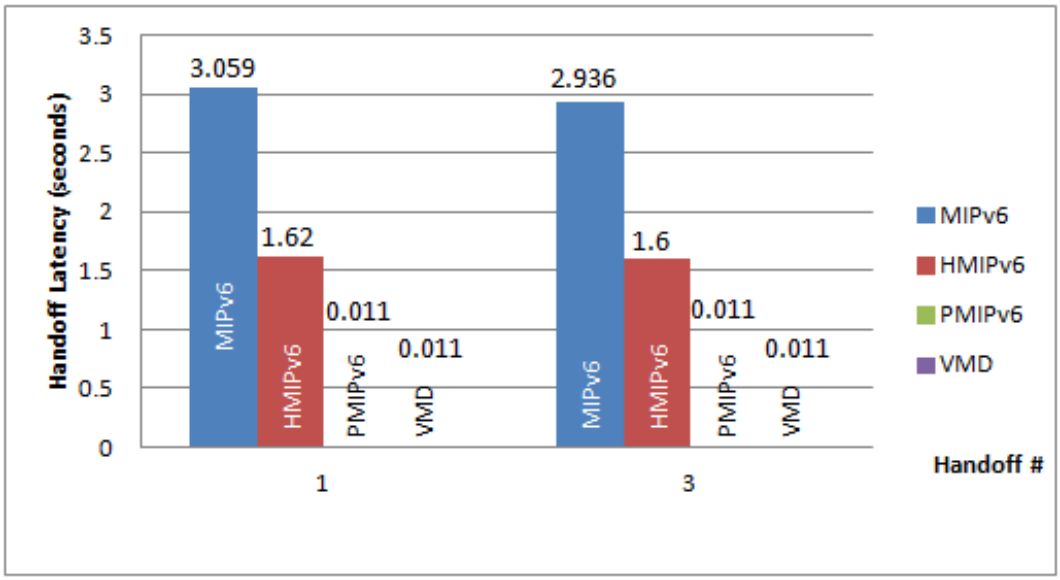


Figure 4.12: Handoff latency results for the tier-2 deployment of HMIPv6, PMIPv6 and VMD, and MIPv6 - observed in OPNET simulations.

4.3.2 Packet Loss

Number of packet loss is presented in Fig. 4.13. The difference between the packet loss results of tier-2 and tier-1 deployment scenarios are proportional to the handoff latency differences between tier-1 and tier-2 deployment scenarios. The change in the packet loss is not significant because data packets transmitted with uniform inter arrival time of 0.01 s.

4.3.3 Signaling Overhead

Fig. 4.14 illustrates the signaling overhead observed for the tier-2 deployment of the protocols. MIPv6 performs the same compared to the results illustrated in Fig. 4.6 as its deployment is not changed. Fig. 4.14 depicts that HMIPv6 and PMIPv6 have reduced signaling overhead results: 224 and 304 bytes, respectively compared to the results illustrated in Fig. 4.6 and belong to tier-1 deployment of the protocols in Section 4.4. The reason of low signaling overhead is that in the current scenario, mobile agents of the protocols are deployed in tier-2 clouds which decreases the number of hops that a mobility message has to travel as compared to the scenario where the protocols deployed in tier-1 cloud in AS. On the other hand, signaling overhead in VMD stays the same since tier-2 cloud handles the mobility utilizing collaborative management in both scenarios.

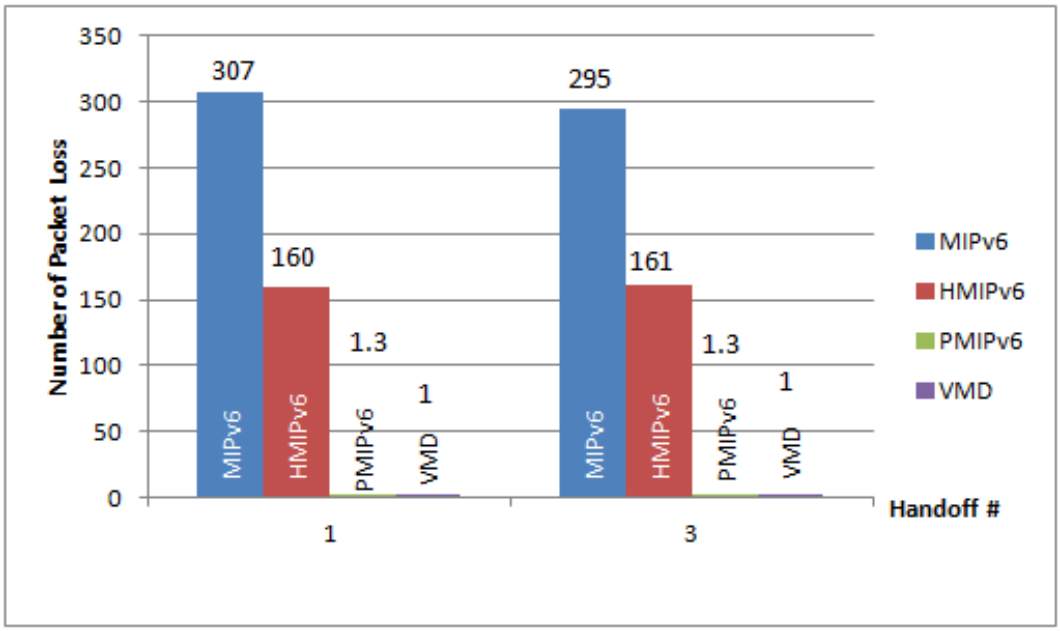


Figure 4.13: Number of packets loss during the handoffs supported by HMIPv6, PMIPv6 and VMD, deployed in tier-2 cloud and MIPv6 - observed in OPNET simulations.

4.4 Multiple-AS Deployment of the Protocols for Inter-AS Roaming

In this section, we present the handoff latency, signaling overhead and packet loss results for handoff performance of the VMD deployed across multiple ASes and ISPs as illustrated in Fig. 4.15. For the comparative analysis MIPv6, HMIPv6, and PMIPv6 are also deployed on the same network in Fig. 4.15. In the simulation, we compare three different scenarios: (i) VMD is deployed at ISP level; (ii) HMIPv6 is deployed in AS 1 and AS 2, and MIPv6 is used for macro-mobility; and (iii) PMIPv6 is deployed in the ASes, and MIPv6 is used for macro-mobility. In each scenario, the mobile node moves with speed of 30 km/h and makes three different handoffs: (i) handoff 1 (intra-cloud), (ii) handoff 2 (inter-cloud), and (iii) handoff 3 (inter-AS) in the network depicted in Fig. 4.15. The recorded values are averaged over 20 simulations with different seeds.

Access routers operate on 802.11g with a data rate of 54 Mbps and send L2 beacons at every 20 ms while router advertisements are uniformly distributed between 0.5 s and 1 s. Coverage areas of neighbor access routers are overlapped. Configuration of all wired nodes are identical and all wired links have data rate of 5 Mbps. To include the effect

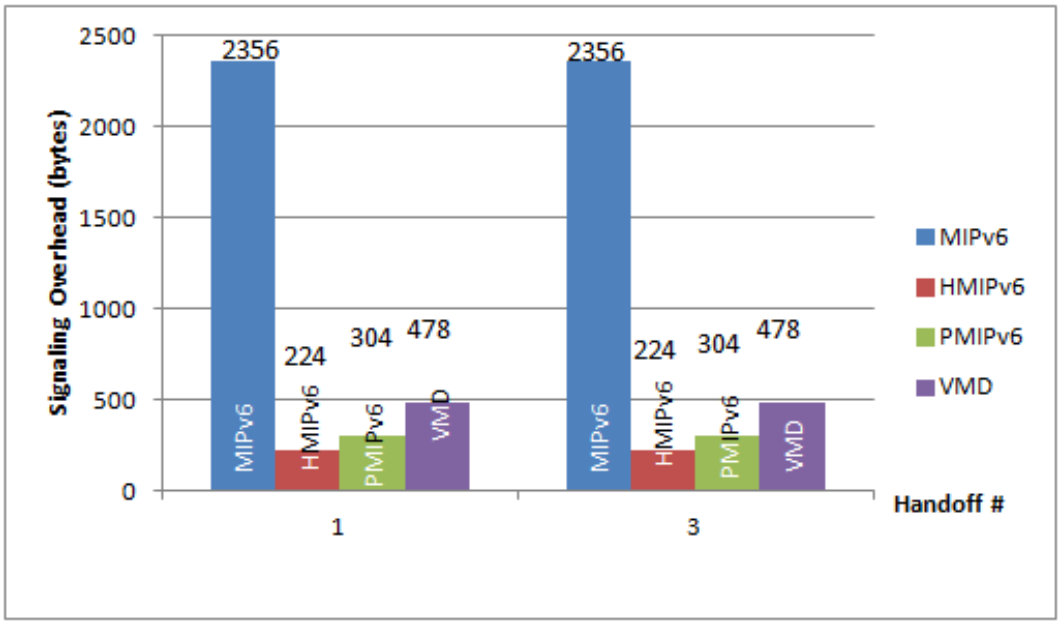


Figure 4.14: Signaling overhead results during the handoffs at the tier-2 deployment of HMIPv6, PMIPv6 and VMD, and MIPv6 - observed in OPNET simulations.

of distances, link delays between access routers and the tier-2 clouds, and root clouds in each AS are 2 ms and 4 ms respectively. The link delays between the ASes, ISP C and the correspondent node are 10 ms. The link delays help including the effect of distances. The data traffic from the correspondent node to the mobile node has uniform packet interarrival time of 10 ms. The network parameter settings are mainly in parallel with the ones in Section 4.2.

4.4.1 Handoff Latency

Fig. 4.16 provides handoff latency incurred by four mobility protocols under three different types of handoff. We confirmed with the simulation logs that each protocol executes all the processes stated in the analytical models in Section 4.1 and handoff latency is composed of only these processes. The only assumption is that mobile nodes are allowed to access a network after AAA procedure is completed. MIPv6 has the highest latency because it is a host-based macro-mobility protocol. As stated in (4.6), the mobile node has to initiate binding update, route optimization with the home agent, the correspondent node, in addition to movement detection and duplicate address detection processes for every handoff. HMIPv6 has reduced latency around 1.6 s for handoff 1 and 2 because the

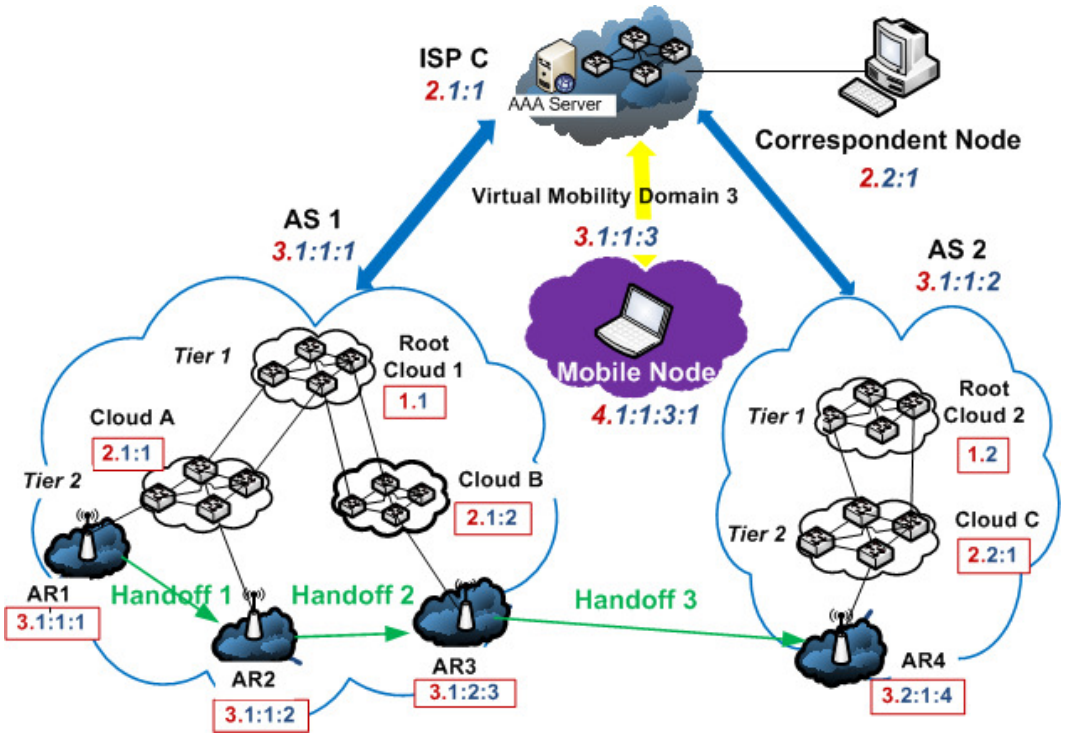


Figure 4.15: Detailed view of the domain 3 in Fig. 6.1 and it is explained in Section 3.4.

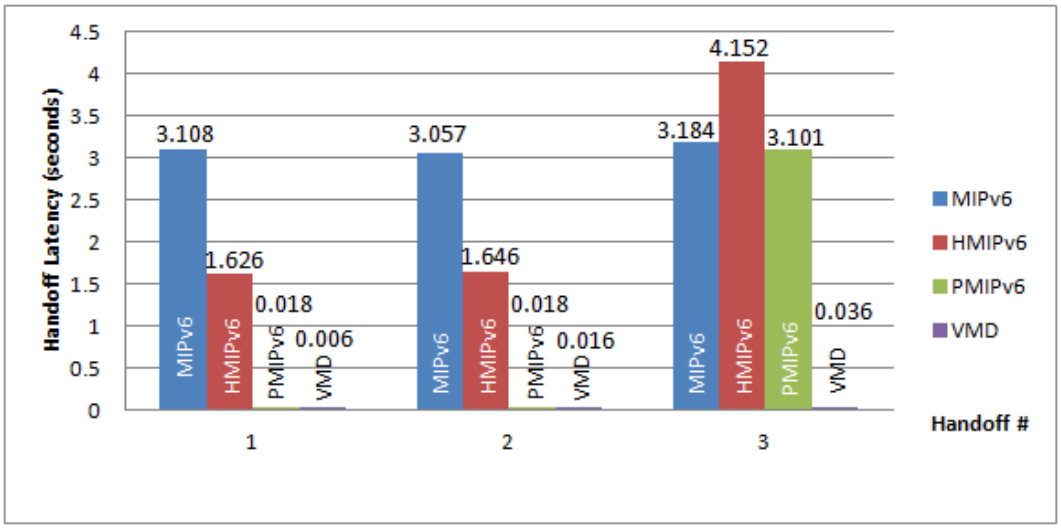


Figure 4.16: Handoff latency results during the three handoff types: (1) intra-cloud, (2) inter-cloud, and (3) inter-AS, observed in OPNET simulations.

mobile node's regional care-of address does not change, and mobility is visible within the AS only. Thus, route optimization and binding update to the correspondent node and to the home agent are avoided as also expressed in (4.14). On the other hand, PMIPv6 and VMD has latency in magnitudes of ms for handoff 1 and 2. That is mainly because these protocols are network-based, and they avoid movement detection [128] and duplicate address detection [137] processes that take around 0.5 s and 1 s, respectively. Further, an address to the mobile node is provided by the network and mobility control messaging only occurs between wired nodes - that also save time. In handoff 1, VMD performs 12 ms better than PMIPv6 because mobility is limited to tier-2 cloud, Cloud A - benefiting from the collaborative management scheme. In handoff 2, PMIPv6 and VMD performs almost same because access routers have to communicate with Root Cloud in both case.

In handoff 3, HMIPv6 and PMIPv6 have to execute MIPv6 as a macro-mobility protocol because the mobile node moves into another mobility domain. Therefore, they have latency of 4 s and 3 s. This difference can be understood from presence/absence of the terms in (4.13), (4.14), (4.17) and (4.18). However, VMD has latency of only 36 ms, because even when the mobile node moves between the ASes, it is still under the same VMD and uses the same address. The control messaging occurs between ISP C, AR3, AR4 and the intermediate clouds on the path to update their proxy AAA servers and forwarding bases.

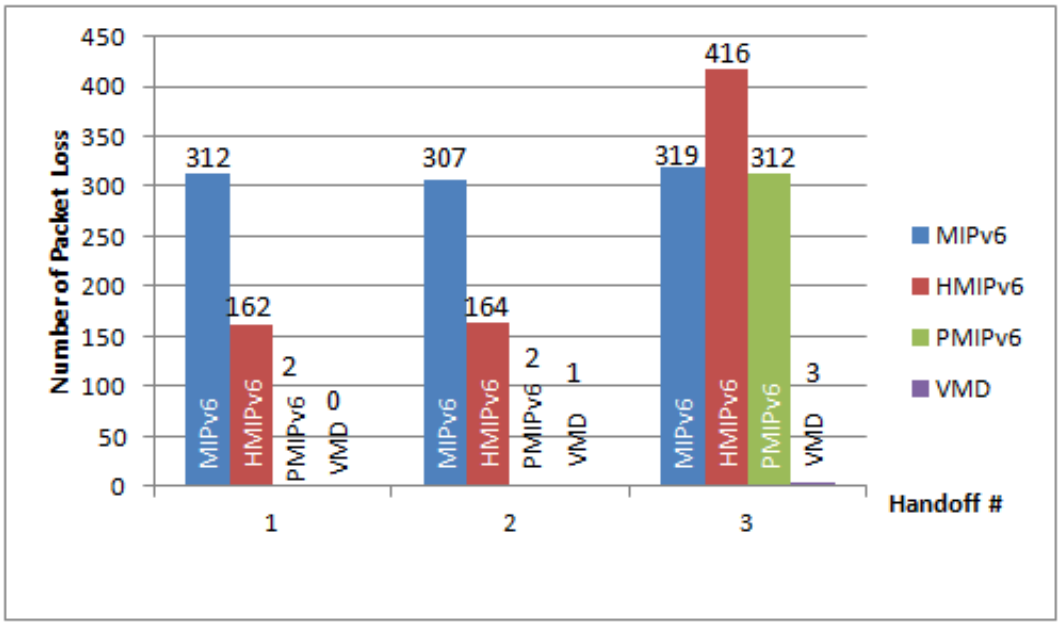


Figure 4.17: Number of packets lost during (1) intra-cloud, (2) inter-cloud, and (3) inter-AS handoffs observed in OPNET simulations.

4.4.2 Packet Loss

In Fig. 4.17, the packet losses during the handoffs are plotted. The results are proportional to the handoff latency because the data traffic from the correspondent node to the mobile node is constant. Routers operate at full-mode hence data packets are delivered at every 10 ms.

4.4.3 Signaling Overhead

Fig. 4.18 is the plot of signaling overhead recorded during handoffs using OPNET simulator. The simulation results match exactly with the numerical results using signaling overhead models in Section 4.1. MIPv6 has very high signaling overhead, 2876 bytes because of the binding message exchange and route optimization occur between the mobile node, the home agent, and the correspondent node as stated in (4.8).

In handoff 1 and 2, the signaling overhead of HMIPv6, PMIPv6 and VMD are less than 610 bytes and the results confirm to (4.8), (4.15), (4.19) and (4.23). In these handoffs, the mobility related signaling is constrained to AS 1 where HMIPv6 and PMIPv6 are deployed, and VMD limits signaling to AS 1 due to the collaborative management scheme. PMIPv6 has higher overhead compared to HMIPv6 because PMIPv6 messages are lengthy

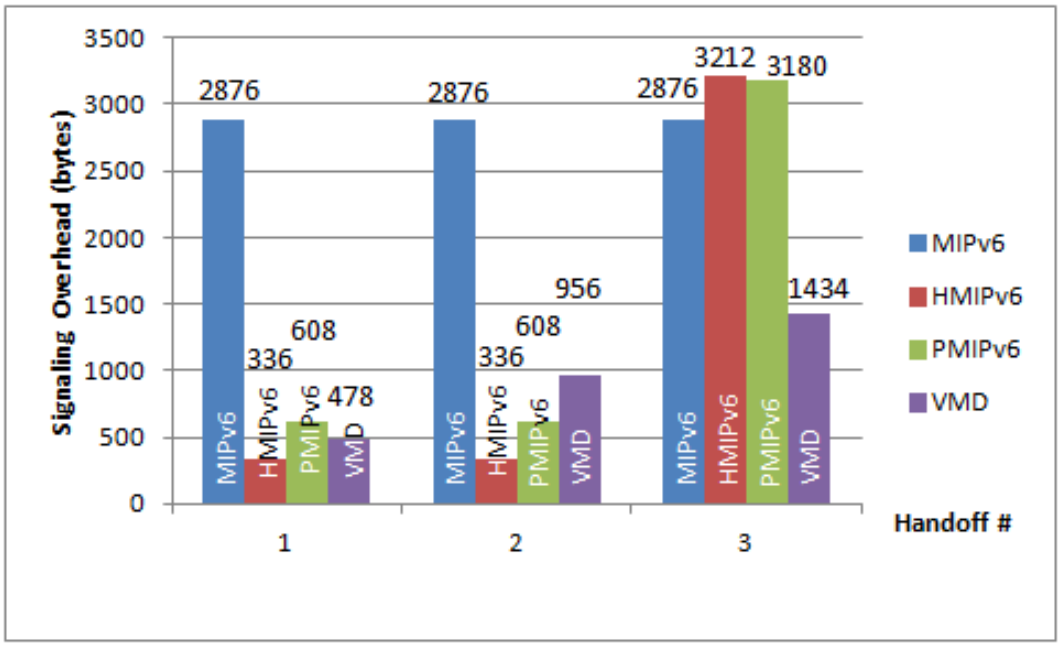


Figure 4.18: Signaling overhead results observed for three handoff types: (1) intra-cloud, (2) inter-cloud, and (3) inter-AS, observed in OPNET simulations.

as presented in Table 4.1. In handoff 3, inter-AS roaming, HMIPv6 and PMIPv6 has increased signaling overhead around 3000 bytes because MIPv6 protocol is executed to provide connectivity with the correspondent node and the home agent. In the case of VMD, the same type of mobility control message exchange happens as in previous handoffs. Except the common anchor cloud is now ISP C, hence the old access router and the new access router communicate with the mobility agent in ISP C. During handoff 3, VMD incurs only 1434 bytes of signaling overhead.

4.5 Summary

In this chapter, we present the analytical models for the handoff latency and signaling overhead in MIPv6, HMIPv6, PMIPv6, and the VMD. Then, we compare these protocols in terms of mobile node identification, ways of handling a handoff, and locations and functions of the different mobility agents. In Section 4.2, we conduct the comparative performance analysis of the protocols, first for the intra-AS roaming scenario mainly over latency, then signaling overhead, and packet loss metrics. We validate the numeric results based on the analytical models with the OPNET simulation-based results. Then, we extend the intra-AS deployment of the protocols to tier-2 of the AS to observe the

effect of the domain size change on the performance of these protocols. Deployment of HMIPv6 and PMIPv6 to a lower tier brings lower latency and less signaling overhead at the expense of mobility domain size as presented in Section 4.3. However, the VMD maintains the same latency and signaling overhead values independent from the tier that it is deployed because of the collaborative management approach. In Section 4.4, we deploy the VMD across multiple ASes to assess its performance during the inter-AS roaming. The handoff latency, packet loss, and signaling overhead results show that the VMD is more scalable for higher tier deployment as compared to IPv6-based mobility protocols.

Table 4.3: VMD performance benchmark results for handoff (H)

VMD vs.	*Latency*						*Signaling Overhead*					
	Sec. 4.2		Sec. 4.3	Sec. 4.4			Sec. 4.2		Sec. 4.3	Sec. 4.4		
	H1&3	H2&4	H1&3	H1	H2	H3	H1&3	H2&4	H1&3	H1	H2	H3
MIPv6	99%	99%	99%	99%	99%	98%	79%	59%	79%	83%	66%	50%
HMIPv6	99%	98%	99%	99%	99%	99%	-42%	-184%	-113%	-42%	-184%	55%
PMIPv6	38%	0%	0%	66%	11%	98%	21%	-57%	-57%	21%	-57 %	54%

Table 4.3 presents the summarization of the handoff latency and the signaling overhead improvements achieved by the VMD as compared to aforementioned IPv6-based protocols. The packet loss results are not included in the table given that they are proportional to the handoff latency results since the data packets are sent from the correspondent to the mobile node with constant inter-arrival time. The results in Table 4.3 were retrieved from the following equation.

$$Performance = \frac{XIPv6 - VMD}{XIPv6} \tag{4.26}$$

where X in $XIPv6$ denotes the first letter in the abbreviation of the protocol, e.g., MIPv6, HMIPv6 and PMIPv6. The protocol names in (4.26) represent either handoff latency or signaling overhead performance for the specified protocol in the OPNET simulation scenarios: (i) the tier-1 deployment scenario for intra-AS roaming presented in Section 4.2, (ii) the tier-2 deployment scenario for intra-AS roaming presented in Section 4.3, and (iii) the multiple-AS deployment scenario for inter-AS roaming presented in Section 4.4. As stated in Section 4.2 and 4.3, handoffs 1 and 3 are intra-cloud handoffs while handoff s2 and 4 are inter-cloud handoffs. In Section 4.4, handoff 1 is an intra-cloud handoff, handoff 2 is an inter-cloud handoff, and handoff 3 is inter-AS handoff. The VMD manages all types of handoffs in approximately 98% less time compared to MIPv6 and HMIPv6, because MIPv6 and HMIPv6 are host-based macro- and micro-mobility protocols, respectively. They require a mobile node to be involved in mobility-control message exchanges with the home agent in MIPv6 and the mobility anchor point in HMIPv6. Compared to

PMIPv6 handoff latency results, the VMD brings improvements in changing magnitudes depending on the handoff type. The VMD and PMIPv6 are both network-based mobility protocols, and they limit mobile node involvement with handoff management. The same handoff latency results are observed in VMD and in PMIPv6 during handoffs 2 and 4 at the tier-1 deployment scenario and handoffs 1 and 3 at the tier-2 deployment scenario, which is due to the fact that locations of local mobility anchor in PMIPv6 and the common anchor cloud in the VMD are the same. However, results observed in tier-1 and tier-2 deployment of the VMD and PMIPv6 imply that deploying the VMD to upper tier results in a larger mobility domain for a user without sacrificing handoff performance while PMIPv6 shows increase in latency. The VMD, deployed across multiple ASes, presented in Section 4.4, outperforms PMIPv6 the most during the inter-AS handoff due to the fact that the VMD can be deployed to upper tiers in the FCT internetworking model and is not restricted to a domain, e.g., AS while PMIPv6 is a micro-mobility protocol, and it needs MIPv6 to handle inter-domain handoff. The handoff latency performance achieved by VMD states that deploying VMD to upper tiers does not cause performance degradation due to the collaborative management scheme introduced by the VMD.

In terms of signaling overhead, the VMD performs better than MIPv6 in changing magnitudes, due to the change of the common anchor cloud in the VMD, depending on the handoff type. Further, MIPv6 forces binding updates, route optimization messaging between mobile node, home agent, and correspondent node for each type of handoff. Compared to PMIPv6, the VMD causes 21% less signaling overhead during intra-cloud handoff at tier-1 deployment and multiple-AS deployment scenarios due to the collaborative management scheme, which avoids communication to the main mobility manager. During the inter-AS handoff, the VMD performs 50%, 55%, and 54% better than MIPv6, HMIPv6, and PMIPv6, respectively, due to the fact that the handoff occurs in the same mobility domain for the VMD while HMIPv6 and PMIPv6 need to collaborate with MIPv6 to handle the handoff. The VMD removes the differentiation of micro- and macro-mobility by being deployed to any tier in the FCT model, which results in the varying mobility domain sizes, allowing a mobile node to use the same address in the mobility domain, and handling each handoff in the same manner using collaborative management scheme leveraging the forwarding bases and the proxy AAA servers.

Chapter 5

Handoff Cost Framework

The advent of new internetworking architectures and associated mobility architectures makes the assessment and comparison of handoff performance difficult. A number of handoff metrics are required to assess the seamless handoff capability of a mobility management scheme, such as registration costs, latency in handoff, data loss during handoff, and signaling overhead that are incurred by the mobile user. This makes comparison of performance with legacy techniques also difficult. To address this concern, these metrics should be assessed in a cohesive manner. The mobile user should be allowed to decide on the best mobility scheme based on his mobility profile and the costs he is willing to incur. Hence, we introduce a new unified handoff assessment metric and handoff cost framework, which accounts for all metrics of interest mentioned above.[1]

In the literature, there are performance comparison studies following non-user centric approaches that focus on network resource usage. The signaling cost incurred in MIPv6 [1], Fast MIPv6 [34], HMIPv6 [35] , and PMIPv6 [38] is analyzed in [42]. Handoff management related messaging, packet delivery, and packet tunneling costs are formulated to find their impact on the network resource consumption. Jong-Hyouk et al. [138] conduct a comparative performance study in terms of the signaling cost of the aforementioned protocols. Kong et al. [39] examine HMIPv6 and PMIPv6 in terms of delay

[*] Portions of this chapter previously appeared as:
H. Tuncer, N. Shenoy, A. Kwasinski, J. F. Hamilton, and S. Mishra, A novel user-centric handoff cost framework applied to the Virtual Mobility Domains and IPv6-based mobility protocols, *Global Telecommunications Conference (GLOBECOM 2012)*, vol. 2578, no. 2584, pp. 3-7, Dec. 2012.
H. Tuncer, A. Kwasinski, and N. Shenoy, Performance Analysis of Virtual Mobility Domain Scheme vs. IPv6 Mobility Protocols, *Elsevier Computer Networks Journal*, Volume 57, Issue 13, 9 September 2013, Pages 2578-2596, ISSN 1389-1286.

[1] In this framework, we do not aim to capture network service usage cost for regular communication or the economics of network pricing. We focus only on the cost of handoff support as perceived by the mobile user.

caused during a handoff. Lee et al. [110] examine the wireless power consumption cost of HMIPv6 and PMIPv6 due to location update and packet delivery. These schemes are discussed in detail in Chapter 2 Section 6.

In this chapter, we present a novel and user-centric handoff cost framework to analyze handoff performance of different mobility schemes. The proposed framework helps examine the impacts of registration costs, signaling overhead, and data loss for the Internet-connected mobile users employing a unified cost metric. The framework is applied to IPv6-based mobility protocols such as HMIPv6 and PMIPv6 to show the framework's flexibility and adaptability. Using the framework, we compare the handoff performance of IPv6-based mobility protocols to the VMD-based protocol. The outcomes indicate firstly the applicability of the handoff cost framework and the unified cost metric in assessing different mobility schemes including IPv6-based mobility protocols. Secondly, using the framework the handoff performance achieved with VMD-based protocol is three and nine times superior to PMIPv6 and HMIPv6, respectively. We further find out the optimal VMD tier that a mobile user should register depending on his/her mobility and cost preferences.

The contribution of this work is threefold:

- We introduce a novel handoff cost framework that can be used to analyze handoff performance of different mobility schemes. The proposed framework helps examine the impacts of registration costs, signaling overhead, and data loss for the Internet-connected mobile users employing a unified cost metric. The framework provides a user-centric approach that allows a mobile user to analyze mobility schemes depending on the mobility preferences, needs, and costs that he is willing to incur.
- The framework can be adopted to assess the performance of different mobility protocols. In this chapter, we illustrate its use in IPv6-based protocols and the VMD-based protocol.
- We apply the proposed handoff cost framework to VMD-based protocol. VMDs overlap and can be deployed to any tier in the proposed Internet architecture as explained in Section 3. This allows the mobile user to register under a VMD in any tier depending on his/her mobility needs and cost consideration. We analyze the effect of each parameter in the handoff cost framework on the total cost and optimum VMD that the mobile user should register.

The rest of this chapter is organized as follows. We introduce the handoff cost framework and then apply it to the VMD-based and IPv6-based protocols in Sections 5.1, 5.3 and 5.4. We provide discussion of the results obtained by applying the proposed framework in Section 5.5. We then present how to find the optimal VMD for a mobile user based on the proposed framework in Section 5.6. We give the concluding remarks in Section 5.7.

5.1 The Handoff Cost Framework

Let $H(d_x)$ denote the handoff cost for a mobile user initially registered with a mobility domain d_x that is managed by protocol x. The framework considers (i) the storage cost ($Sto(d_x)$) at proxy AAA servers and routing/forwarding entities; (ii) the signaling cost ($Sig(d_x)$), which is incurred due to mobility control message signaling to support handoff; and (iii) the cost of data loss ($Data\_Loss(d_x)$) due to handoff latency. All the cost components are expressed in bytes. $H(d_x)$ can thus be defined by:

$$H(d_x) = w_p \cdot (\mu \cdot Sto(d_x) + \theta \cdot Sig(d_x)) + w_d \cdot Data\_Loss(d_x), \tag{5.1}$$

where w_d, w_p, μ, and θ are the weights. The aim of introducing these weights is that each handoff cost component (signaling, storage, and data loss cost) may have a different impact on the total handoff cost depending on mobile user preferences, network settings, or service provider requirements. These weights are introduced to identify the relative impact of each cost component on the total handoff cost. We retrieve these weights from the mobile user or the service provider.

We recognize a mobile user's relative sensitivity to the cost of data loss, compared to signaling and storage costs, with w_d. The mobile user's relative sensitivity to the storage and the signaling costs that are incurred on the network is denoted with w_p. The summation of w_p and w_d is given a weight of one. The values for these weights can be changed depending on the application properties and the mobile user's preferences. If losing a connectivity due to a handoff is very costly for a mobile user, then we expect w_d to receive higher values. In the case that storage and signaling overhead is not costly for a mobile user, then we expect the w_p gets values closer to zero.

Further, we differentiate the cost of storage and signaling by introducing weights μ and θ. To illustrate this, storage may not be costly compared to the signaling overhead, depending on the current technology and the service provider. In that case, we expect the value of μ to be lower than the value of θ. The summation of μ and θ is one.

The mobility of a mobile user is managed by mobility agents (MAs) in a mobility domain. The handoff cost framework considers the costs incurred during the activities of mobility agents, specifically signaling cost and data loss cost. Therefore, it is important to identify the mobility agents that are involved in handoff-related activities. Here we will introduce sets $A_{in}(d_x)$ and $A_{out}(d_x)$ to identify the mobility agents that handle the mobility of a mobile user. Let $A_{in}(d_x)$ be the set of all mobility agents in the domain d_x while $A_{out}(d_x)$ be the set of all mobility agents that are not in the domain d_x. We express these sets as

follows:

$$A_{in}(d_x) = \{MA \mid MA \in d_x\} \text{ and}$$
$$A_{out}(d_x) = \{MA \mid MA \notin d_x\}. \tag{5.2}$$

Not all of these mobility agents in these aforementioned sets handle the mobility of a mobile user. Let $B_{in}(d_x)$ denote the subset of mobility agents in $A_{in}(d_x)$ that actually handle the in-domain handoffs of a given mobile user. Likewise, let $B_{out}(d_x)$ denote the subset of mobility agents in $A_{out}(d_x)$ that actually handle the out-of-domain handoffs of the mobile user. Mathematically, these sets can be defined as

$$B_{in}(d_x) = \{MA \in A_{in}(d_x) \mid HO(MA) > 0\},$$
$$B_{out}(d_x) = \{MA \in A_{out}(d_x) \mid HO(MA) > 0\}, \tag{5.3}$$

where $HO(MA)$ denotes the number of the handoffs that are handled by a mobility agent. In-domain and out-of-domain handoffs depend on the coverage area of the initially registered mobility domain d_x.

We will present storage, signaling, and data loss costs in the next sections.

5.1.1 Storage Cost

User's registration cost to a mobility domain d_x is represented as the storage cost because the mobile user-related information has to be stored at proxy AAA servers and routing/forwarding tables by the service provider to support mobility. Further, the mobile user needs to register with a different mobility domain temporarily if he moves out of d_x to which he is initially registered. The total storage cost is thus given by

$$Sto(d_x) = Sto_{in}(d_x) + Sto_{out}(d_x), \tag{5.4}$$

which is the summation of the storage cost at the initially registered domain ($Sto_{in}(d_x)$) and the temporarily registered domains out-of d_x ($Sto_{out}(d_x)$).

5.1.2 Signaling Cost

Signaling cost is the sum of the signaling overheads incurred during the mobile user's registration to a domain and during handoffs. Signaling overhead is calculated by multiplying the mobility-control message sizes by the number of hops that each message

travels.[2]

Mobility-control messages are sent during the initial registration and during in-domain and out-of-domain handoffs. The signaling cost consists of the initial registration cost ($Sig_{init}(d_x)$); the in-domain signaling cost ($Sig_{in}(d_x)$), due to all the in-domain handoffs; and the out-of-domain signaling cost ($Sig_{out}(d_x)$), due to all the out-of-domain handoffs. The total signaling cost is thus given by

$$Sig(d_x) = Sig_{init}(d_x) + Sig_{in}(d_x) + Sig_{out}(d_x). \tag{5.5}$$

The in-domain signaling cost is the summation of the signaling overhead during each handoff that is handled by mobility agents in d_x. It is defined by

$$Sig_{in}(d_x) = \sum_{MA \in B_{in}(d_x)} HO(MA) \cdot C_{in}(MA), \tag{5.6}$$

where $HO(MA)$ is the number of the handoffs handled by the mobility agent MA, which is a member of $B_{in}(d_x)$ in the equation, and ($C_{in}(MA)$) denotes the signaling overhead that is incurred due to a handoff handled by MA.

The out-of-domain signaling cost is the summation of the cost of signaling required for each out-of-domain handoff that is handled by mobility agents out of d_x. It is defined by

$$Sig_{out}(d_x) = \tau \cdot \sum_{MA \in B_{out}(d_x)} HO(MA) \cdot C_{out}(MA), \tag{5.7}$$

where ($C_{out}(MA)$) denotes the signaling overhead that is incurred due to an out-of-domain handoff handled by MA. τ is an external service cost multiplier represents the extra cost of getting services from a service provider that serves a user *temporarily* as the user roams into its service area. This happens when a mobile user moves out of the coverage area of the initially registered mobility domain d_x. The exact value of parameter τ depends on the business relationship between the service provider with which the mobile user is registered permanently and the service provider with which the mobile user is registered temporarily. The general constraints for τ are

$$\tau_{min} \leq \tau \leq \tau_{max}, \tag{5.8}$$

where τ_{min} and τ_{max} denote the minimum and the maximum values that τ can have, respectively. The cost of getting service from the temporarily registered service provider could be more than the cost of the service received from the permanently registered service provider as explained in Section 5.5.

[2]Unit of both signaling cost and signaling overhead is byte.

5.1.3 Data Loss Cost

$Data\_Loss(d_x)$ denotes the cost of the mobile user's data loss due to handoff latency. $Data\_Loss(d_x)$ is

$$Data\_Loss(d_x) = Data\_Loss_{in}(d_x) + Data\_Loss_{out}(d_x), \tag{5.9}$$

where $Data\_Loss_{in}(d_x)$ is data loss cost due to in-domain handoffs while $Data\_Loss_{out}(d_x)$ is data loss cost due to out-of-domain handoffs.

The cost of data loss due to in-domain handoffs is defined by

$$Data\_Loss_{in}(d_x) = \lambda_s \cdot R_{data} \cdot \sum_{MA \in B_{in}(d_x)} HO(MA) \cdot D_{in}(MA), \tag{5.10}$$

where λ_s is the number of active communication sessions that the mobile node maintains per unit time, and R_{data} denotes the average number of data bytes per session. $D_{in}(MA)$ denotes the handoff latency that is incurred due to the mobile user's handoff that is handled by MA, which is member of $B_{in}(d_x)$ in the equation.

The cost of data loss due to out-of-domain handoffs is defined by

$$Data\_Loss_{out}(d_x) = \lambda_s \cdot R_{data} \cdot \sum_{MA \in B_{out}(d_x)} HO(MA) \cdot D_{out}(MA), \tag{5.11}$$

where $D_{out}(MA)$ denotes the handoff latency during a mobile user's handoff out of the coverage area of the initially registered domain d_x. The handoff is handled by MA , which is a member of $B_{out}(d_x)$, which denotes the set of mobility agents that handle the out-of-domain handoffs.

5.2 Application to VMD

In this section, we will present the application of the handoff cost framework to VMD. We draw the Floating Cloud Tiered internetworking model in Fig. 5.1 where the ISPs and ASes are at tier 1 to 6.[3] Let the VMD that a mobile user is initially registered be d_{vmd}. In the application of the handoff cost framework to a VMD, we will use all the equations in Section 5.1 that are Eqns. (5.1 - 5.9) by replacing d_x with d_{vmd} as the domain is managed by VMD. Instead of re-writing the equations that are presented previously, we will provide the formulation of each cost component in the aforementioned equations.

[3]We omitted tier 3 and 4 not to clutter the figure. They are shown as dots in the figure.

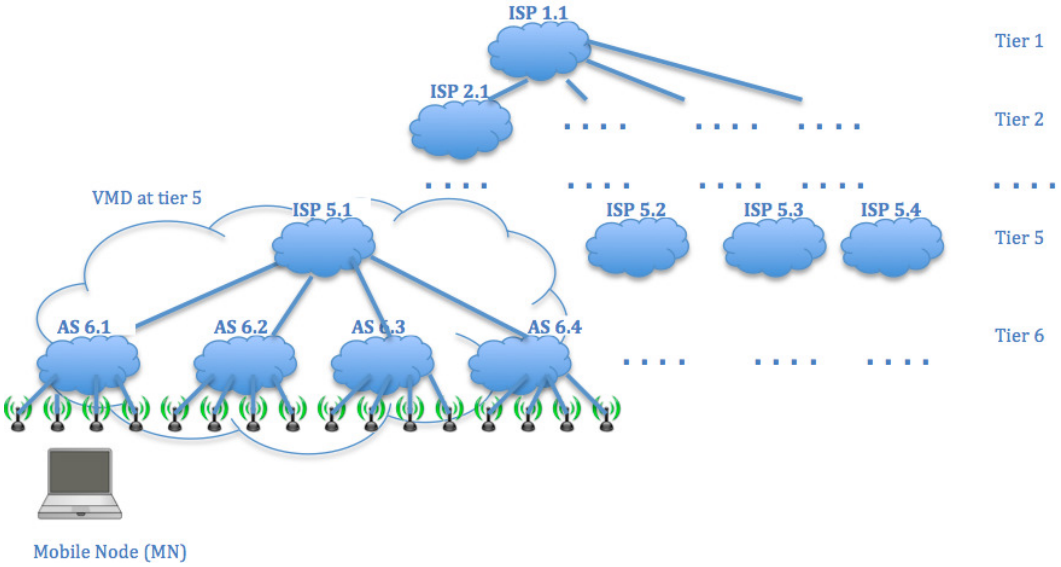


Figure 5.1: The Floating Cloud Tiered internetworking model.

5.2.1 Storage cost components at VMD

Rewriting Eqn. (5.4) by replacing d_x with d_{wmd}, $Sto_{in}(d_{vmd})$ denotes the cost of storing the mobile node profile at all the proxy AAA servers ($Sto_{AAA_{in}}(d_{vmd})$), and the forwarding bases ($Sto_{FB_{in}}(d_{vmd})$) in the initially registered VMD, d_{vmd}. $Sto_{in}(d_{vmd})$ is defined by

$$Sto_{in}(d_{vmd}) = Sto_{AAA_{in}}(d_{vmd}) + Sto_{FB_{in}}(d_{vmd}). \tag{5.12}$$

The cost of storing the mobile node profile at all the proxy AAA servers in d_{vmd} is defined by

$$Sto_{AAA_{in}}(d_{vmd}) = N[AAA_{in}(d_{vmd})] \cdot \eta, \tag{5.13}$$

where $AAA_{in}(d_{vmd})$ is set of proxy AAA servers in the VMD d_{vmd} that is defined by

$$AAA_{in}(d_{vmd}) = \{AAA \mid AAA \in d_{vmd}\}; \tag{5.14}$$

$N[AAA_{in}(d_{vmd})]$ is the number of proxy AAA servers in $AAA_{in}(d_{vmd})$; and η denotes the cost of storing a mobile node profile at a proxy AAA server in d_{vmd}. The unit of η is bytes.

So far the equations above are valid for any topology such as balanced tree or unbalanced tree. The formulation of $N[AAA_{in}(d_{vmd})]$ in Eqn. (5.15) below is specific to the

topology that VMD is deployed to, which is balanced tree topology [139] in Fig. 5.1. In a balanced tree topology, each node has exactly the same number of children nodes, γ and the bottom tier has a value of K. Therefore, $N[AAA_{in}(d_{vmd})]$ is defined by

$$N[AAA_{in}(d_{vmd})] = \sum_{i=T(d_{vmd})}^{K-1} \gamma^{i-T(d_{vmd})}, \tag{5.15}$$

where $T(d_{vmd})$ returns the tier value of the root of d_{vmd}. To illustrate, if d_{vmd} is the VMD that is rooted at ISP 5.1 in Fig. 5.1, then $T(d_{vmd})$ will return 5.

The forwarding bases are also located at each node in the topologythe same as proxy AAA servers. The storage cost in forwarding bases ($Sto_{FB_{in}}(d_{vmd})$ in Eqn. (5.12)) is defined by

$$Sto_{FB_{in}}(d_{vmd}) = N[FB_{in}(d_{vmd})] \cdot \delta, \tag{5.16}$$

where $FB_{in}(d_{vmd})$ is a set that consists of forwarding bases in the VMD d_{vmd}, that is defined by

$$FB_{in}(d_{vmd}) = \{FB \mid FB \in d_{vmd}\}; \tag{5.17}$$

$N[FB_{in}(d_{vmd})]$ is the number of forwarding bases in $FB_{in}(d_{vmd})$; and δ denotes the cost of storing forwarding information at a forwarding base. The unit of δ is bytes. In our tiered topology, $N[FB_{in}(d_{vmd})]$ is formulated by

$$N[FB_{in}(d_{vmd})] = \sum_{i=T(d_{vmd})}^{K} \gamma^{i-T(d_{vmd})}. \tag{5.18}$$

A mobile node may move out of the initially registered d_{vmd} and this movement causes the mobile node to make an out-of-domain handoff and register with another VMD temporarily. The mobility agents in $B_{out}(d_{vmd})$ handles mobility. To retrieve the VMDs to which the mobile node is temporarily registered, we introduce a function called V. This function accepts $B_{out}(d_{vmd})$, the set of mobility agents that handle the mobile node's handoff to out-of initially registered d_{vmd}. V returns a set of VMDs to which the mobile node is temporarily registered. Let d_{vmd-t} denote the temporarily registered VMD.

The temporary registration to a VMD results in a storage cost that is referred as $Sto_{out}(d_{vmd})$ in Eqn. (5.4). The storage cost in the temporarily registered VMD consists of the cost for storing a mobile node profile temporarily at the proxy AAA servers ($Sto_{AAA_{out}}(d_{vmd})$) and forwarding bases ($Sto_{FB_{out}}(d_{vmd})$). $Sto_{out}(d_{vmd})$ is defined by

$$Sto_{out}(d_{vmd}) = \tau \cdot \alpha \cdot [Sto_{AAA_{out}}(d_{vmd}) + Sto_{FB_{out}}(d_{vmd})], \tag{5.19}$$

where α is the fraction of time spent out of the initially registered VMD, d_{vmd} and τ is the cost multiplier for the extra cost of getting services temporarily from any other mobility domain.

The formulation of $Sto_{AAA_{out}}(d_{vmd})$ and $Sto_{FB_{out}}(d_{vmd})$ is given in 5.20 and 5.24, respectively. The formulation is based on two cases between d_{vmd} and d_{vmd-t} as follows:

- *Case 1:* d_{vmd} and d_{vmd-t} do not overlap and do not have any common anchor clouds.
- *Case 2:* d_{vmd-t} completely contains d_{vmd} .

The cost of storing a mobile node profile temporarily at the proxy AAA servers ($Sto_{AAA_{out}}(d_{vmd})$ in Eqn. (5.19)) is defined by

$$Sto_{AAA_{out}}(d_{vmd}) = \sum_{d_{vmd-t} \in V(B_{out}(d_{vmd}))} N[AAA_{out}(d_{vmd}, d_{vmd-t})] \cdot \eta, \tag{5.20}$$

where $AAA_{out}(d_{vmd}, d_{vmd-t})$ denotes the set of proxy AAA servers that reside only in d_{vmd-t}. $AAA_{out}(d_{vmd}, d_{vmd-t})$ is defined as

$$AAA_{out}(d_{vmd}, d_{vmd-t}) = AAA_{in}(d_{vmd-t}) \setminus AAA_{in}(d_{vmd}) = \{AAA \in d_{vmd-t} \mid AAA \notin d_{vmd}\}. \tag{5.21}$$

$N[AAA_{out}(d_{vmd}, d_{vmd-t})]$ is the number of proxy AAA servers in $AAA_{out}(d_{vmd}, d_{vmd-t})$ and is defined by

$$N[AAA_{out}(d_{vmd}, d_{vmd-t})] = \begin{cases} N[AAA_{in}(d_{vmd-t})] & (Case\ 1) \\ N[AAA_{in}(d_{vmd-t})] - N[AAA_{in}(d_{vmd})] & (Case\ 2), \end{cases} \tag{5.22}$$

In our illustrative case, where VMD is deployed on a balanced tree topology, $N[AAA_{out}(d_{vmd}, d_{vmd-t})]$ is calculated by

$$N[AAA_{out}(d_{vmd}, d_{vmd-t})] = \begin{cases} \sum_{i=T(d_{vmd-t})}^{K-1} \gamma^{i-T(d_{vmd-t})} & (Case\ 1) \\ \sum_{i=T(d_{vmd-t})}^{K-1} \gamma^{i-T(d_{vmd-t})} - \sum_{i=T(d_{vmd})}^{K-1} \gamma^{i-T(d_{vmd})} & (Case\ 2). \end{cases} \tag{5.23}$$

Above, we simply count the number of proxy AAA servers in d_{vmd-t}. The tier value of the root of d_{vmd-t} is found by $T(d_{vmd-t})$. We start counting from tier $T(d_{vmd-t})$ down to the bottom of the balanced tree. We extract the number of nodes in d_{vmd} as needed in Case 2.

The storage cost in the forwarding bases in d_{vmd-t} ($Sto_{FB_{out}}(d_{vmd})$ in Eqn. (5.19)) is defined by

$$Sto_{FB_{out}}(d_{vmd}) = \sum_{d_{vmd-t} \in V(B_{out}(d_{vmd}))} N[FB_{out}(d_{vmd}, d_{vmd-t})] \cdot \delta, \tag{5.24}$$

where $FB_{out}(d_{vmd}, d_{vmd-t})$ is set of forwarding bases that reside only in d_{vmd-t}. $FB_{out}(d_{vmd}, d_{vmd-t})$ is defined by

$$FB_{out}(d_{vmd}, d_{vmd-t}) = FB_{in}(d_{vmd-t}) \setminus FB_{in}(d_{vmd}) = \{FB \in d_{vmd-t} \mid FB \notin d_{vmd}\}. \tag{5.25}$$

$N[FB_{out}(d_{vmd}, d_{vmd-t})]$ is the number of forwarding bases in $FB_{out}(d_{vmd}, d_{vmd-t})$. It is defined by

$$N[FB_{out}(d_{vmd}, d_{vmd-t})] = \begin{cases} N[FB_{in}(d_{vmd-t})] & (Case\ 1) \\ N[FB_{in}(d_{vmd-t})] - N[FB_{in}(d_{vmd})] & (Case\ 2). \end{cases} \tag{5.26}$$

In our illustrative case, where VMD is deployed on a balanced tree topology, $N[FB_{out}(d_{vmd}, d_{vmd-t})]$ is calculated by

$$N[FB_{out}(d_{vmd}, d_{vmd-t})] = \begin{cases} \sum_{i=T(d_{vmd-t})}^{K} \gamma^{i-T(d_{vmd-t})} & (Case\ 1) \\ \sum_{i=T(d_{vmd-t})}^{K} \gamma^{i-T(d_{vmd-t})} - \sum_{i=T(d_{vmd})}^{K} \gamma^{i-T(d_{vmd})} & (Case\ 2), \end{cases} \tag{5.27}$$

where we did a simple arithmetic calculation of the number of the forwarding bases that will handle out-of-domain handoffs in d_{vmd-t}. Forwarding bases reside at each cloud in d_{vmd-t}.

5.2.2 Signaling cost components at VMD

We will provide the formulation of the cost components that are introduced in Eqn. (5.5). $Sig_{init}(d_{vmd})$ denotes the signaling overhead that is incurred during the initial registration to d_{vmd}. $Sig_{init}(d_{vmd})$ is defined by

$$Sig_{init}(d_{vmd}) = 2 \cdot (K + 1 - T(d_{vmd})) \cdot m, \tag{5.28}$$

where m is the size of a mobility-control message as presented in Table 4.1. The reasoning for these values are given in Section 4.1.2. The total number of hops that a message travels is $2 \cdot (K + 1 - T(d_{vmd}))$, including the wireless link. Please refer to Section 3 for details of the path that the handoff-control message follows. K is the number of tiers in the topology.

$C_{in}(MA)$ in Eqn. (5.6) denotes the signaling overhead that is incurred due to the mobile node's handoff that is handled by the mobility agent in the common anchor cloud in the VMD. It is defined by

$$C_{in}(MA) = 3 \cdot (K - T(MA)) \cdot m, \tag{5.29}$$

where $T(MA)$ returns the tier value of the mobility agent, and $3 \cdot (K - T(MA))$ denotes the total number of links that the message follows. The mobile node is connected to the access points at the tier K, hence $K - T(MA)$ denotes the number of links between MA and the access point. The mobility messages travel three times in total for connection request, approval, and acknowledgement. Please refer to Section 3 for details of the path that the handoff-control message follows.

$C_{out}(MA)$ in Eqn. (5.7) denotes the signaling overhead that is incurred during the mobile node's handoff to the coverage area outside of the initially registered VMD. These handoffs are handled by the mobility agent in $B_{out}(d_{vmd})$. There will be two different cases for the mobility agent:

- *Case 1:* The first time that the mobility agent in the common anchor cloud is handling the mobile node's handoff, the mobility agent does not have the mobile node's profile, and it needs to retrieve it from the domain with which the mobile node was previously registered. Note that the mobile node's previous domain is assumed to be one hop away from the common anchor cloud where the mobility agent is residing.
- *Case 2:* The mobility agent has the mobile node's profile information due to previously supporting the mobile node's handoff.

$C_{out}(MA)$ is thus given by

$$C_{out}(MA) = \begin{cases} 3 \cdot (K - T(MA) + 1) \cdot m & (Case\ 1) \\ 3 \cdot (K - T(MA)) \cdot m & (Case\ 2). \end{cases} \tag{5.30}$$

As expressed above, in *Case 1*, a mobility-control message travels more hops due to the need for retrieving the mobility profile from the initially registered VMD. Please refer to Section 3 for details of the path that a handoff-control message follows.

5.2.3 Data loss cost components at VMD

We will present the formulation of the cost components that are introduced in Eqns. (5.9 - 5.11). $D_{in}(MA)$ in 5.10 denotes the handoff latency that is incurred due to the mobile node's handoff, which is handled by the mobility agent in $B_{in}(v)$. $D_{in}(MA)$ is defined by

$$D_{in}(MA) = 2 \cdot (K - T(MA)) \cdot t_w, \tag{5.31}$$

where $2 \cdot (K - T(MA))$ denotes the number of hops that a mobility-control message with m bytes has to travel, as explained in Section 3. A one-way transmission delay between two wired nodes that are one hop away is denoted with t_w and it is defined by

$$t_w = \frac{m}{B_w} + L_w + p_q, \tag{5.32}$$

where m is the mobility-control message size, B_w is the bandwidth of a wired link, L_w is the propagation delay of the wired link [125], and p_q is the average processing and queuing time of a packet at a router [129].

$D_{out}(MA)$ in 5.11 denotes the handoff latency during a mobile node's handoff out of the coverage area of the initially registered VMD. The handoff is handled by MA, which is a member of $B_{out}(v)$, which denotes the set of MA that handle the out-of-domain handoffs. $D_{out}(MA)$ is defined by

$$D_{out}(MA) = \begin{cases} 2 \cdot t_{wl} + 2 \cdot (K - T(MA) + 1) \cdot t_w & (Case\ 1) \\ 2 \cdot (K - T(MA)) \cdot t_w & (Case\ 2), \end{cases} \tag{5.33}$$

where t_{wl} is a one-way wireless transmission delay between the mobile node and an access router. t_{wl} is defined by

$$t_{wl} = \frac{m}{B_{wl}} + L_{wl} + p_q, \tag{5.34}$$

where m is the mobility-control message size, B_{wl} is the bandwidth of a wireless link, and L_{wl} is the propagation delay of the wireless link [125].

The numerical values assigned to the aforementioned parameters can be found in Tables 4.1 and 5.1. The numerical analysis is provided in Section 5.5.

5.3 Application to HMIPv6

This section presents the application of the handoff cost framework, that is presented in Section 5.1, to Hierarchical IPv6 (HMIPv6). HMIPv6 is a micro-mobility protocol that handles the mobility of a mobile user within a mobility domain. See Section 2.2.2 for details. Note that we assume that mobile user is already registered to a home network and macro-mobility of user is handled by MIPv6 (See Section 2.2.1 for details). We will apply the handoff cost framework to HMIPv6 by replacing d_x with d_{hmipv6} and MA with MAP in Eqns. (5.1 - 5.9) as the mobility agent in HMIPv6 is called mobility anchor point (MAP). Instead of re-writing the equations that are presented previously, we will provide the formulation of each cost component.

A mobile user may go out of the home network that is managed by MIPv6 and move to a domain that is managed by HMIPv6. Each time a mobile node registers with a HMIPv6 domain or moves to a new HMIPv6 domain, following processes are executed and they cause signaling overhead and handoff delay. For example, binding updates by the mobile node with the home agent, the mobility anchor point and the correspondent node create signaling overheads, denoted with $C_{BU}(MN, HA)$, $C_{BU}(MN, MAP)$, and $C_{BU}(MN, CN)$, respectively. Signalling overhead due to route optimization process between the home agent and the correspondent node is denoted by C_{RO}. The notation for the delays during the binding update processes follows the format of $T_{BU}(X, Y)$ as stated in Section 4.1.1. A handoff delay due to route optimization is denoted by T_{RO}. A handoff in the HMIPv6 domain further requires movement detection (MD); duplicate address detection (DAD); and authentication, authorization, and accounting (AAA) of the mobile user and related handoff delays are denoted with T_{MD}, T_{DAD}, and T_{AAA}, respectively. See Chapter 2.2 for details of how MIPv6 and HMIPv6 handle handoff. We intentionally do not provide a detailed derivation of signaling, storage, or handoff delay here as we provided them in Section 4.1. The numerical values for the parameters in the following cost equations can be found in Table 5.1. The numerical analysis is provided in Section 5.5.

5.3.1 Storage cost components at HMIPv6

The storage cost consists of the cost of storing the mobile node profile and routing information at the permanently registered domain and the domain that the mobile node is temporarily registered.

The initial storage cost consists of the storage cost at proxy AAA servers and routers in d_{hmipv6}, that are denoted by $Sto_{AAA_{in}}(d_{hmipv6})$ and $Sto_{R_{in}}(d_{hmipv6})$, respectively. We formulate the initial storage cost as follows:

$$Sto_{in}(d_{hmipv6}) = Sto_{AAA_{in}}(d_{hmipv6}) + Sto_{R_{in}}(d_{hmipv6}). \tag{5.35}$$

The storage cost at a temporarily registered domain that is out of d_{hmipv6} is defined by

$$Sto_{out}(d_{hmipv6}) = \tau \cdot \alpha \cdot (Sto_{AAA_{out}}(d_{hmipv6}) + Sto_{R_{out}}(d_{hmipv6})), \tag{5.36}$$

where $Sto_{AAA_{out}}(d_{hmipv6})$ denotes the storage cost at proxy AAA servers out of the initially registered domain while $Sto_{R_{out}}(d_{hmipv6})$ denotes the storage cost at routers out of the initially registered domain.

5.3.2 Signaling cost components at HMIPv6

When mobile user moves out of home network and registers to an HMIPv6 domain, there needs to be a binding update with the home agent in the home network, correspondent

node, and mobility anchor point in addition to the route optimization process.[4] Therefore, the signaling cost due to initial registration in HMIPv6 is defined by [110][5]

$$Sig_{init}(d_{hmipv6}) = C_{BU}(MN, HA) + C_{RO} + C_{BU}(MN, CN) + C_{BU}(MN, MAP). \quad (5.37)$$

The signaling overhead due to an in-domain handoff is defined by:

$$C_{in}(MAP) = C_{BU}(MN, MAP). \quad (5.38)$$

As generically expressed in Eqn. (5.6), multiplication of $C_{in}(MAP)$ with number of in-domain handoffs gives signaling cost due to in-domain handoffs.

The signaling overhead due to an handoff to the temporary domain is defined by:

$$C_{out}(MAP) = C_{BU}(MN, HA) + C_{RO} + C_{BU}(MN, CN) + C_{BU}(MN, MAP). \quad (5.39)$$

where MAP belongs to a temporarily registered domain. $C_{out}(MAP)$ is part of the out-of-domain signaling cost calculation that is formulated generically in Eqn. (5.6).

The numerical values assigned to the aforementioned parameters can be found in Table 4.1. The numerical analysis is provided in Section 5.5.

5.3.3 Data loss cost components at HMIPv6

The data loss cost in HMIPv6 is calculated using Eqn. (5.9). Here, we will only provide in-domain and out of domain handoff delay formulations as these are specific to HMIPv6.

In-HMIPv6-domain handoff delay that comprises movement detection delay (T_{MD}), duplicate address detection delay (T_{DAD}), authentication delay (T_{AAA}), and binding update delay ($T_{BU}(MN - MAP)$) is defined by [39]

$$D_{in}(MAP) = T_{MD} + T_{DAD} + T_{AAA} + T_{BU}(MN, MAP). \quad (5.40)$$

Out-of-HMIPv6-domain handoff delay comprises additional delays because of MIPv6 processes, such as home agent binding update delay ($T_{BU}(MN, HA)$), route optimization delay (T_{RO}), and correspondent node binding update delay ($T_{BU}(MN, CN)$) [39]. Thus

$$\begin{aligned} D_{out}(MAP) =& T_{MD} + 3 \cdot T_{DAD} + T_{AAA} + T_{BU}(MN, MAP) \\ &+ T_{BU}(MN, HA) + T_{RO} + T_{BU}(MN, CN). \end{aligned} \quad (5.41)$$

T_{DAD} is multiplied by three to include duplicate address detection delay for the on-link care-of address at the mobile node and the regional care-of address at the mobility anchor point and at the home agent. The definition and the detailed formulation of the parameters above can be found in Section 4.1.

[4] See Section 2.2.1 for details of route optimization.

[5] Note that we assume the mobile user is already registered to a home network and now he moves out of the home network.

5.4 Application to PMIPv6

This section presents the application of the handoff cost framework presented in Section 5.1 to Proxy Mobile IPv6 (PMIPv6). PMIPv6 is a micro-mobility protocol that handles the mobility of a mobile user within a mobility domain. See Section 2.2.3 for details. We assume that the mobile user is already registered to a home network that is managed by MIPv6. We will apply the handoff cost framework to PMIPv6 by replacing d_x with d_{pmipv6} and *MA* with *LMA* in equations 5.1 - 5.9 as a mobility agent in PMIPv6 is a local mobility anchor (LMA). Instead of re-writing the equations that are presented previously, we will provide the formulation of each cost component.

During the mobile node's handoff in the PMIPv6 domain, binding update messaging between a mobility access gateway (MAG) and a local mobility anchor occurs. Each time a mobile node moves out of a home network and registers to the PMIPv6 domain, binding update messaging between mobile node and home agent and correspondent node occur in addition to route optimization process. Further, authentication, authorization, and accounting of the mobile node and also duplicate address detection processes are executed. These processes cause signaling overhead and delay. We intentionally do not provide a detailed derivation of signaling, storage, or handoff delay as we provided them in Section 4.1 along with the already existing literature. The numerical values for the parameters in the following cost equations can be found in Table 5.1. The numerical analysis is provided in Section 5.5.

5.4.1 Storage cost components at PMIPv6

The storage cost consists of the cost of storing the mobile node profile and routing information at the initially registered domain and the domain to which the mobile node will be temporarily registered.

The initial storage cost consists of the storage cost at proxy AAA servers and routers in d_{pmipv6}, which are denoted by $Sto_{AAA_{in}}(d_{pmipv6})$ and $Sto_{R_{in}}(d_{pmipv6})$, respectively. Storage cost at the initially registered domain is

$$Sto_{in}(d_{pmipv6}) = Sto_{AAA_{in}}(d_{pmipv6}) + Sto_{R_{in}}(d_{pmipv6}). \tag{5.42}$$

The storage cost at the temporarily registered domain that is out of d_{pmipv6} is

$$Sto_{out}(d_{pmipv6}) = \tau \cdot \alpha \cdot (Sto_{AAA_{out}}(d_{pmipv6}) + Sto_{R_{out}}(d_{pmipv6})). \tag{5.43}$$

where $Sto_{AAA_{out}}(d_{pmipv6})$ denotes the storage cost at proxy AAA servers out of d_{pmipv6} and $Sto_{R_{out}}(d_{pmipv6})$ denotes the storage cost at routers that are out of the domain of d_{pmipv6}.

5.4.2 Signaling cost components at PMIPv6

When mobile node moves out of the home network and registers to a PMIPv6 domain, PMIPv6 performs a binding update with the local mobility anchor and mobility access gateway ($C_{BU}(MAG, LMA)$) after MIPv6 messaging is executed. The signaling cost due to initial registration is defined by

$$Sig_{init}(d_{pmipv6}) = C_{BU}(MN, HA) + C_{RO} + C_{BU}(MN, CN) + C_{BU}(MAG, LMA) \tag{5.44}$$

in [110].

The signaling overhead due to an in-domain handoff is

$$C_{in}(LMA) = C_{BU}(MAG, LMA). \tag{5.45}$$

The signaling cost due to in-domain handoffs is multiplication of number of in-domain handoffs with $C_{in}(LMA)$, as expressed in Eqn. (5.6).

The signaling overhead due to an out-of-domain handoff to a temporary domain is

$$C_{out}(LMA) = C_{BU}(MN, HA) + C_{RO} + C_{BU}(MN, CN) + C_{BU}(MAG, LMA) \tag{5.46}$$

because when a mobile node goes out of the PMIPv6 domain, it needs to register with another domain, which requires the same steps as in the initial registration. The signaling cost due to out-of-domain handoffs is multiplication of number of out-of-domain handoffs with $C_{out}(LMA)$, as expressed in Eqn. (5.7).

5.4.3 Data loss cost components at PMIPv6

The data loss cost in PMIPv6 is calculated using Eqn. (5.9). Here, we will only provide in-domain and out-of-domain handoff delay formulations as these are specific to PMIPv6.

In-PMIPv6-domain handoff delay is

$$D_{in}(LMA) = T_{AAA} + T_{BU}(MAG, LMA), \tag{5.47}$$

where $T_{BU}(MAG, LMA)$ denotes the delay due to the binding update with the local mobility anchor and T_{AAA} denotes the delay due to the communication with AAA servers to authenticate the mobile node [39].

When the mobile node initiates an out-of-domain handoff, MIPv6 processes are executed in addition to PMIPv6 as explained in [39]. Hence,

$$D_{out}(LMA) = T_{AAA} + T_{BU}(MAG, LMA) + T_{DAD} + T_{BU}(MN, HA) + T_{RO} + T_{BU}(MN, CN), \quad (5.48)$$

that is summation of a delay due to AAA ($_{AAA}$), delay due to a binding update between the mobility access gateway and the local mobility anchor ($T_{BU}(MAG, LMA)$), delay due to duplicate address detection (T_{DAD}), delay due to a binding update between home agent and mobile node ($T_{BU}(MN, HA)$), delay due to route optimization (T_{RO}) and delay due to a binding update between mobile node and correspondent node ($T_{BU}(MN, CN)$).

5.5 Analytical Results and Discussion

In this section, we present the analytical results based on the equations derived in Sections 5.1, 5.3 and 5.4. In Table 5.1, we list the values chosen to conduct the numerical study. We assign a value for w_p that is lower than the value of w_d to represent the notion that the user is more concerned about data loss due to handoff latency rather than signaling overhead and the storage costs that are incurred on the network. We make sure that the summation of w_p and w_d is 1, because we introduce these weights to differentiate the relative costs. We assign 0.5 to μ and θ, because we do not aim to study the differences between the storage and the signaling costs that are incurred on the system.

We assign 1.1 to τ, considering that the cost of getting service from a temporary service provider is higher than from the permanent service provider. We do not assign a significantly higher value to τ to prevent the costs incurred on the external service provider to dominate all the costs on the initially registered service provider. We assign 6 to K because there are six tiers in our current tiered Internet architecture as identified in [2]. We set γ to 4 to have a symmetric distribution of the clouds on the deployment area. For the remaining parameters, such as mobility control message sizes and bandwidth, we chose the numerical values based on [2, 140–142] and on empirically realistic values. The size of the mobility control messages are presented in Table 4.1.

We assume that a mobile user spends most of his/her time in specific locations (such as home, school, work, and shopping areas, etc.), and he rarely goes to different locations at great distances (such as an overseas vacation). We assume that the repeatedly visited places are in the center of the mobile user's roaming area, while the rarely visited places are close to the edges of the roaming area. A new mobility model based on the centered movement of a mobile user is proposed later in Section 6.1. However, here, we aim to create a mobile user profile to provide a value for $HO(MA)$, which is the number of

handoffs that are handled by the mobility agent. We assume that the mobile node makes 90% of handoffs within AS 6.1, which is the initially registered domain, and the rest of the handoffs are in AS 6.2, AS 6.3, and AS 6.4 in Fig. 5.1. The out-of-domain rate of 10% is chosen as a basis here; however, we have extended our study to the various values of the out-of-domain rate in Section 5.6.1. As a starting point, we assume that the mobile user makes 100 handoffs. However, we have also studied the impact of various numbers of handoffs on handoff cost in Section 5.6.3.

We would like to emphasize that the aim of this numerical study is to demonstrate that our proposed handoff cost framework can be applied for the performance comparison of VMD, HMIPv6, and PMIPv6. Therefore, we use the same parameter values for all the protocols that we compare. The handoff cost framework will continue to operate in the way that it is designed regardless of the values assigned to the parameters. The handoff cost framework will allow combining the aforementioned cost components and provide the total handoff cost that occurs in VMD, HMIPv6, or PMIPv6.

Table 5.1: Numerical values for the handoff cost framework parameters

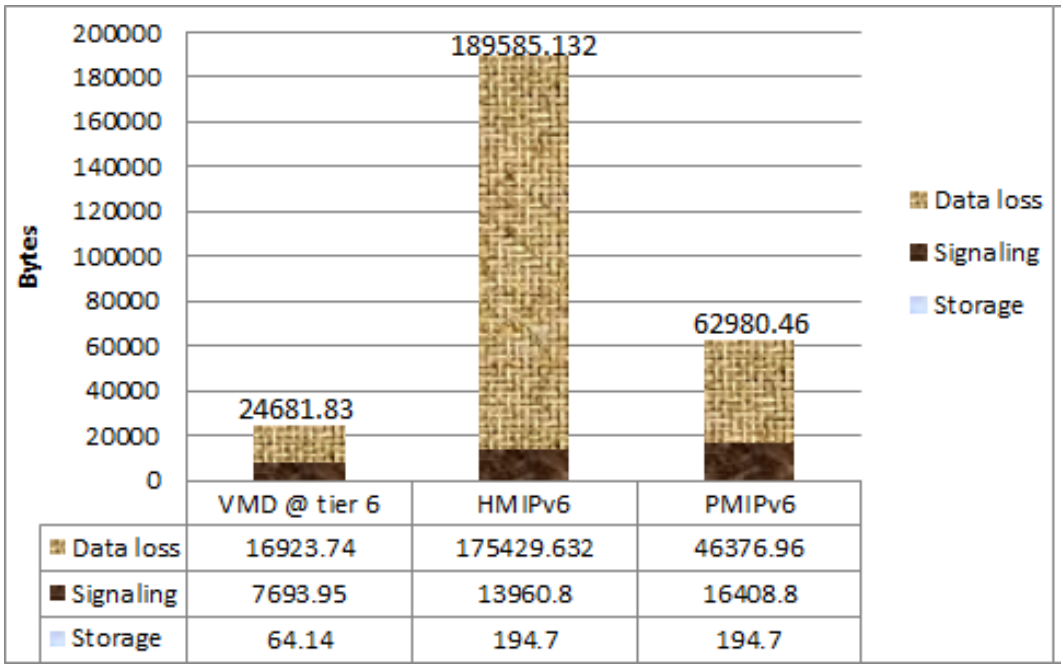
w_p	w_d	K	λ_s	R_{data}	η	L_w	δ
0.3	0.7	7	0.01	100 kB	250 B	0.01 s	10 B
μ	θ	γ	τ	B_{wl}	B_w	L_{wl}	p_q
0.5	0.5	4	1.1	100 Mbps	1 Gbps	0.002 s	0.1 s

5.5.1 The Framework Applied to IPv6-Based Mobility Protocols

This subsection shows that the proposed framework is capable of providing a unified platform for assessing the performance of the VMD, HMIPv6, and PMIPv6 from the mobile user's perspective. The results are presented in Figure 5.2. The network that the VMD is deployed in is AS 6.1 in Figure 5.1. HMIPv6 is deployed in AS 6.1, AS 6.2, AS 6.3, and AS 6.4, while ISP 5.1 can be considered as the mobile node's home network, where MIPv6 handles the inter-domain mobility. PMIPv6 is deployed in AS 6.1, AS 6.2, AS 6.3, and AS 6.4, while ISP 5.1 is the mobile node's home network. The correspondent node is assumed to be in ISP 5.1. We chose these deployment scenarios because we think they are empirically realistic and do not cause any advantage to any protocol among others. We further would like to be consistent with the network setup in Sections 3 and 4.

Figure 5.2 shows the storage, signaling, and data loss costs for the mobile user's handoffs in the VMD, HMIPv6, and PMIPv6. The value at the top of each column in Figure 5.2 is the total cost of handoffs calculated using Eqn. (5.1). For each protocol, the storage cost is the lowest cost (i.e., 64 and 194 bytes) compared to signaling and data loss costs that

are around tens of thousands of bytes. The reason of the storage cost being the lowest is that the number of network nodes is few and the mobile node-related data is smaller than the total mobility control messages. The signaling cost depends on the number of hops that each control message has to travel and the message size. Data loss costs at VMD at tier 6, HMIPv6, and PMIPv6 domains (that are 16923, 175429, and 46376 bytes, relatively) dominate the total handoff cost mainly due to the mobile user's data session size and the importance that he gives to the data loss cost (w_d). The mobile user's data usage value is aligned with the value we used in our numerical studies in Section 4. As we explained in Section 5.1, a mobile user gives more importance to data loss compared to the signaling overhead; hence, the impact of w_d is as expected.



	VMD @ tier 6	HMIPv6	PMIPv6
Data loss	16923.74	175429.632	46376.96
Signaling	7693.95	13960.8	16408.8
Storage	64.14	194.7	194.7

Figure 5.2: Handoff costs for VMD, HMIPv6, and PMIPv6.

In terms of the total handoff cost, HMIPv6 has the highest value because it is a host-based mobility protocol. According to Eqns. (5.38 and 5.40), the mobile node has to initiate mobility control messaging, movement detection, and duplicate-address detection processes during a handoff. PMIPv6 and VMD costs are less than HMIPv6 mainly because these protocols are network based, and they avoid movement detection and duplicate-address detection processes. The VMD has the lowest overall handoff cost because it applies the collaborative mobility management scheme via common anchor clouds; therefore, all of the mobility control messages do not have to go through each

node. Further, the mobile node registers with a larger domain only when it needs to, which keeps storage, signaling, and data loss costs low.

5.5.2 The Framework Applied to VMD

The strength of the framework developed in Section 5.1 is that it accepts the user-related inputs and enables the mobile user to analyze the trends on the different cost components and their impact on the overall handoff cost. Fig. 5.3 provides the individual cost components and the total handoff cost as a function of the initially registered VMD tier.[6] The mobility profile of the mobile user stays the same. The storage cost decreases with increasing tier value mainly because of the decrease in the number of the proxy AAA servers and the forwarding bases that have the mobile node data as given in Eqn. (5.15). Out-of-domain storage costs are incurred only when the mobile node is registered to the VMD at tier 6 because 10% of the mobile node's handoffs are out-of the VMD at tier 6. The signaling cost increases with the decreasing VMD tier, because the control messages for initial registration are transmitted over more hops when the mobile node registers with a VMD at an upper tier. When the mobile node is registered to the VMD at tier 6, the total signaling cost gets to the highest value because 10% of the handoffs will be out of the domain and will cause extra signaling. The extra signaling also causes more delays and proportionally more data loss cost, as expressed in Eqn. (5.9), which explains the peak of the data loss cost for the case of the VMD at tier 6. For the other cases, data loss cost does not change, because all of the handoffs occur within the domain, and they are handled via the collaborative mobility management scheme.

In Fig. 5.3, it can be observed that the storage, signaling, and data loss costs do not change linearly with respect to VMD tier; hence, $H(d_{vmd})$ might be a convex function. The characteristic of $H(d_{vmd})$ motivates us to conduct an optimization study, considering the various values for the parameters in the handoff cost framework, which are mobile user cost sensitivities, mobility profiles, and data communication characteristics.

5.6 Analyzing VMD Performance

Mobile users may have varying mobility preferences and expectations from a mobility protocol. The proposed framework enables a mobile user to apply his/her preferences such as relative sensitivity to the costs imposed by the service provider (w_p), relative sensitivity to the data loss cost (w_d), amount of data usage ($\lambda \cdot R_{data}$), his/her mobility profile with number of handoffs ($HO(MA)$), and the external service cost multiplier (τ).

[6] In the rest of the chapter, the decimal points of the cost values are dropped to provide a better figure presentation and for easier readabilityconsidering that the numbers are large enough to discard the decimal points.

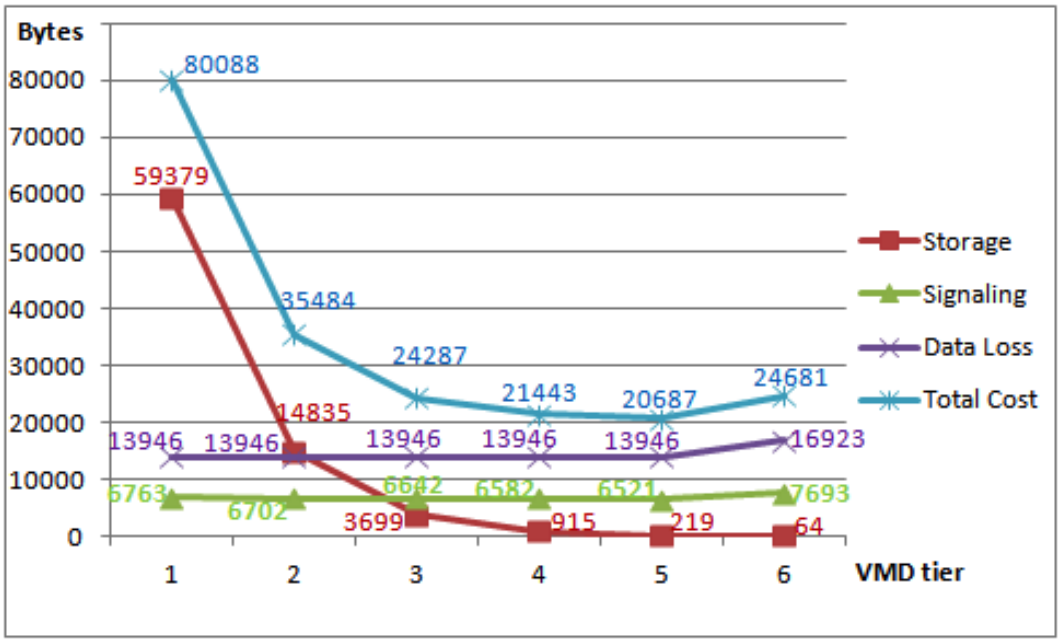


Figure 5.3: Cost vs. initially registered VMDs at tiers 1-6.

[7] In this section, we aim to analyze the effect of each parameter on the handoff cost components. We will gradually change the values of w_p, w_d $\lambda \cdot R_{data}$, $HO(MA)$, and τ one at a time to observe their impact. In this way, we will be creating a set of parameters that can represent the set of different user characteristics.

5.6.1 Effect of Out-of-Domain Handoff Rate

In this subsection, we aim to see the effect of out-of-domain handoff rate on the storage cost, signaling cost, data loss cost, and the optimal VMD tier with which a mobile user should register. We are illustrating our case using the mobile user who makes most of the handoffs in AS 6.1. We assume that the mobility agents in AS 6.2, AS 6.3, AS 6.4, and ISP 5.1 handle the equal number of the out-of-domain handoffs. To accomplish our goal, we vary the out-of-domain rate. Out-of-domain handoff rate means the out-of-domain handoff rate compared to all of the mobile user's handoffs. Fig. 5.4, 5.5 and 5.6 represent the storage cost, signaling cost, and data loss cost, respectively, as a function of the out-of-domain rate for the three different cases, where (i) the mobile node is initially registered with the VMD at tier 6, deployed in AS 6.1; (ii) the mobile node is initially registered with the VMD at tier 5, deployed in ISP 5.1; and (iii) the mobile node is initially registered with

[7] Detailed definition of terms were provided in Section 5.1.

the VMD at tier 4, deployed in the ISP that is the provider of ISP 5.1.[8] We limit our study to these three cases because, for this mobile user, going beyond the VMD at tier 4 does not bring any new insight to the effect of the out-of-domain handoff rate. We limit the out-of-domain handoff rate to 0.2 as the trend of cost values do not change at the higher rates.

The parameter values stated in Table 5.1 and the network setup in Fig. 5.1 are maintained. The external service-cost multiplier is τ, which denotes the cost of temporarily getting service from another service provider maintained at 1.1. We do not assign high values because those high values will significantly minimize the impact of the costs due to the handoffs handled at the initially registered VMD. We would like to observe the impact of the costs handled at the initially registered VMD. However, we also present the effect of varying values of τ in Section 5.6.6.

Fig. 5.4 illustrates the storage cost for a mobile user who is initially registered with the VMD at tier 6. The handoff cost increases 4.78 bytes for every 0.025 increment in the out-of-domain rate.[9] The reason for this steady increase is due to the fraction of time that is spent out of the initially registered VMD, which is denoted by α in Eqn. (5.19). In the cases that the mobile user is registered with the VMD at tier 5 and tier 4, the mobile user has a storage cost of 219 bytes and 915 bytes, respectively, for all out-of-domain handoff rates. In these cases, all of the handoffs are handled by an initially registered VMD, which means that there is not a temporary storage cost. Therefore, the storage costs stay the same for the cases that the mobile user is initially registered with the VMD at tier 5 and tier 4.

Fig. 5.5 depicts the effects of the out-of-domain handoff rate on the signaling cost. The mobile user incurs the lowest storage cost (7290 bytes) in the case that he is initially registered with the VMD at tier 6 where he does not create any out-of-domain handoffs. If the mobile user's out-of-domain handoff rate reaches 0.025, then the signaling cost at the cases of the VMD at tier 6 and tier 5 are the same. Therefore, the mobile user should register with the VMD at tier 5 to have a broader roaming area, which spans ISP 5.1 and all the ASes, which are customers of ISP 5.1. For a 0.025 step up on the out-of-domain handoff rate, the signaling cost increases 100 bytes, 44 bytes, and 44 bytes in the cases of the VMD at tier 6, tier 5, and tier 4, respectively. The reason for the increase in all the cases is that the number of the handoffs handled by the mobility agent in the common anchor cloud at tier 5, which is ISP 5.1, increases proportionally with the increasing out-of-domain handoff rate. In the case that the mobile user is initially registered at VMD at tier 6, there is a steeper increase in cost compared to the other cases that the mobile user

[8] In Fig. 5.1, we omitted ISPs at tier 4 and 5 not to clutter the figure.

[9] Having the bytes in decimal points is due to the weights in the handoff cost framework.

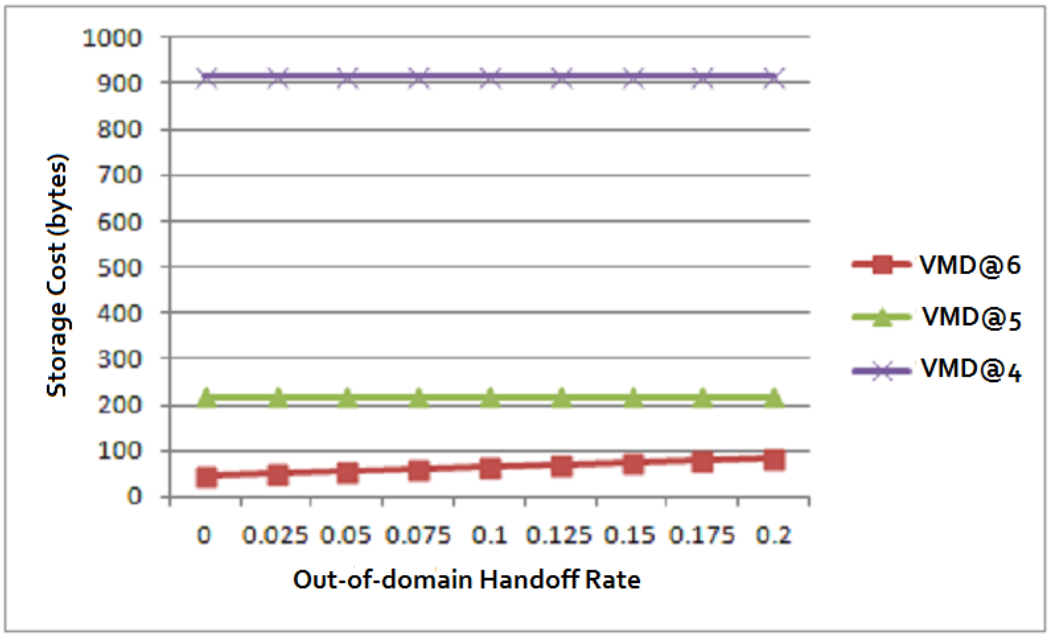


Figure 5.4: Storage cost as a function of the out-of-domain handoff rate.

is registered to the VMDs at tier 4 and 5. The reason for this is that the mobile user has to make more out-of-domain handoffs compared to the VMDs at other tiers, and out-of-domain handoffs are handled by temporary service providers, as explained in Eqn. (5.5). If the mobile user registers with the VMD at tier 5, he will continue having the lowest signaling cost for out-of-domain handoff rates that are more than 0.025 because most of the handoffs are handled by in the VMD at tier 5.

Fig. 5.6 shows the data loss cost depending on the out-of-domain handoff rate for the initially registered VMDs at tier 6, 5, and 4. If all of the mobile user's handoffs happen within the AS 6.1, which means zero out-of-domain handoff rate, then the cost of data loss is the same for all of the initially registered VMD cases. In this case, all the handoffs handled by the mobility agent in AS 6.1 are within the VMDs and, hence, all the handoffs incurred the same delay as expressed in Eqn. (5.9). For each 0.025 increase in the out-of-domain handoff rate, the data loss cost increases 288 bytes, 98 bytes, and 98 bytes in the cases of the VMDs at tier 6, tier 5, and tier 4, respectively. The reason for these increases in all cases is that the number of handoffs handled by the mobility agent in the common anchor cloud at tier 5 increases proportionally with the increasing of the out-of-domain handoff rate. The reason for the higher increase in the case of the VMD at tier 6 is that the out-of-VMD handoffs incur higher latency compared to the cases that all of the handoffs

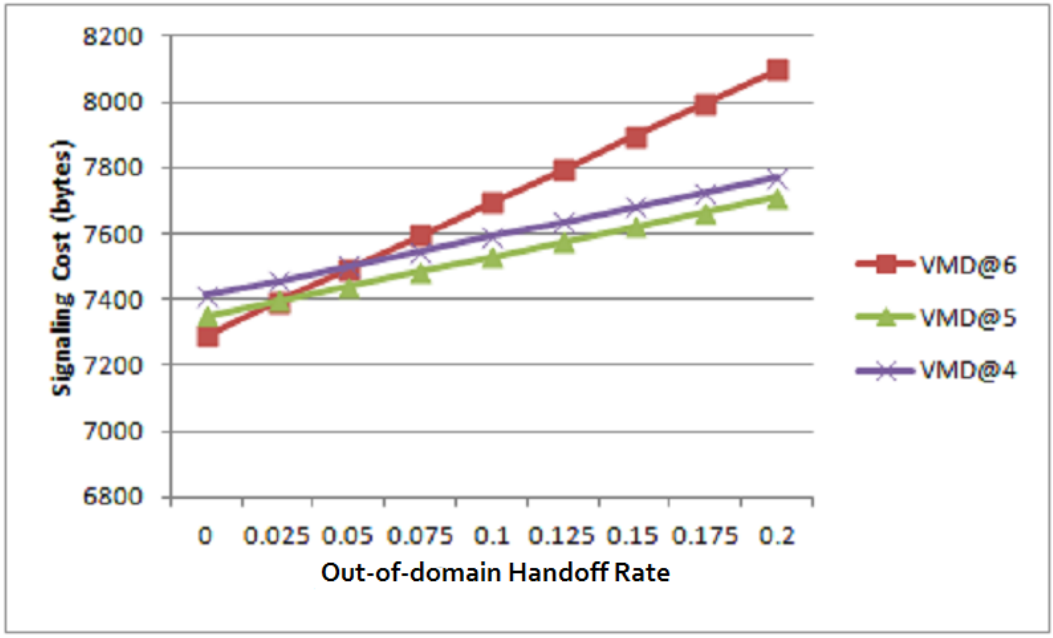


Figure 5.5: Signaling cost vs. the out-of-domain handoff rate for the initially registered VMDs at varying tiers.

are in VMD if the mobile user registered with the VMD at tier 5 or tier 4. Considering only the cost of data loss, the mobile user should register with the VMD at tier 5 or tier 6. The decision of the mobile user in registering between the VMD at tier 6 and the VMD at tier 5 can depend on the storage and signaling costs that are incurred in these cases.

Fig. 5.7 illustrates the total handoff cost for the cases registered with the VMD at tiers 6, 5, and 4 for the mobile user. The total handoff costs are calculated using Eqn. (5.1), which is the sum of the storage, signaling, and data loss costs including the mobile user's sensitivity to these costs. If the mobile user roams within the AS 6.1 all of the time, which means that the out-of-domain handoff rate is 0, then the optimum VMD that the mobile user should register with is the VMD at tier 6 with the lowest total handoff cost (23105 bytes). However, the total handoff cost for the cases of the VMDs at tier 6 and tier 5 are the same (23499 bytes) for a 0.025 out-of-domain rate, and being registered with the VMD at tier 5 brings the lowest cost compared to the other initially registered VMD cases. Therefore, the mobile user should register with the VMD at tier 5 if he does more than, or equal to, 0.025 out-of-domain handoff rate.

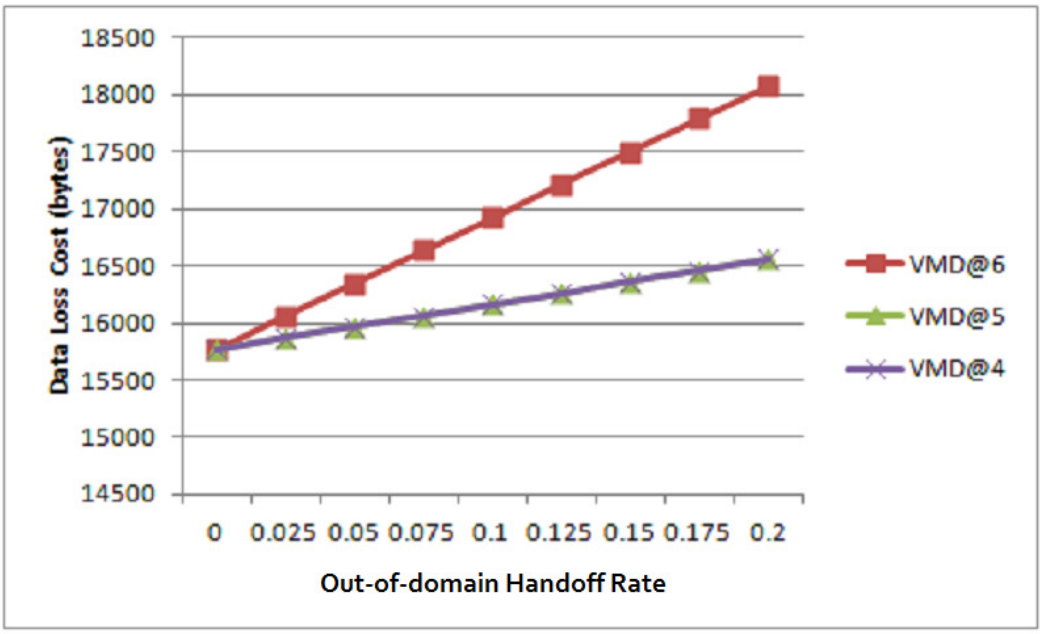


Figure 5.6: Data loss cost as a function of the out-of-domain handoff rate for the initially registered VMDs at varying tiers.

5.6.2 Effect of a Mobile User's Roaming Scope

In this subsection, we aim to find the optimal VMD that a mobile user should register with based on his/her mobility profile. For that purpose, we create four different mobile user profiles. We also introduce a new terminology called mobility reference tier. The mobility reference tier of a mobile user is the tier of the domain where a mobile user roams the most.

- *User_6* makes 80% of handoffs within the AS 6.1 domain, illustrated in Fig 5.1. AS 6.1 is rooted at tier 6 in the topology. Therefore, the mobility reference tier for User_6 is tier 6. User_6 makes the remaining handoffs, which account for 20% of the total number of handoffs, outside of the AS 6.1 domain. These handoffs are considered out of domain. User's out-of-domain handoff rate is 0.2.

- *User_5* makes 80% of handoffs within the ISP 5.1 domain, which is rooted at tier 5 in the topology. Therefore, the mobility reference tier of User_5 is tier 5. User_5's out-of-domain handoff rate is 0.2. The 20% of the handoffs are within the domain supported by service providers rooted at tier 4.

- *User_4* makes 80% of handoffs within the domain rooted at tier 4 in Fig. 5.1. There-

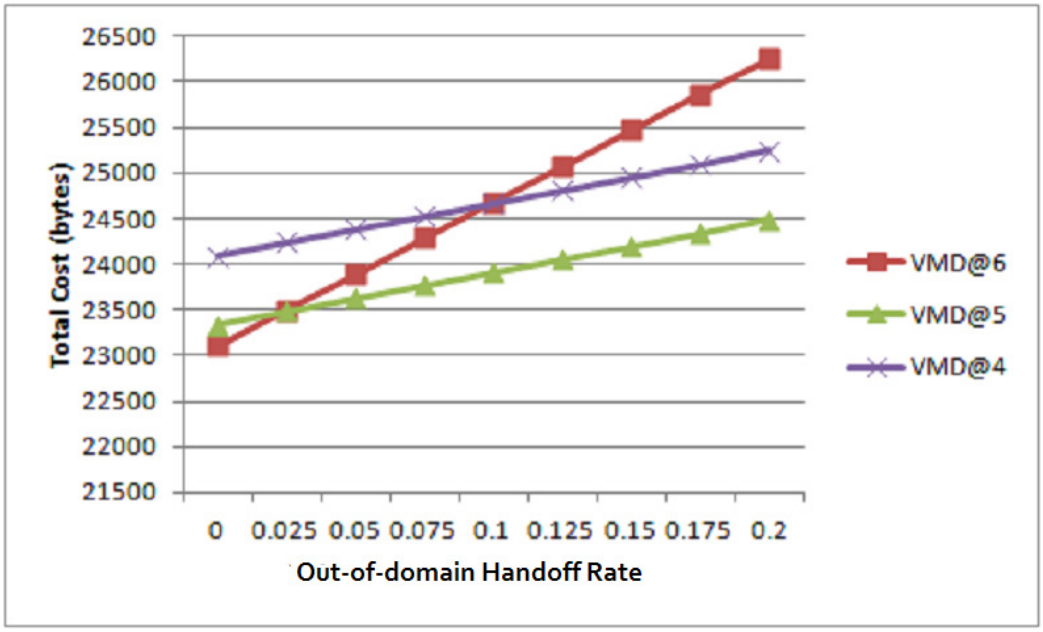


Figure 5.7: Total handoff cost vs. the out-of-domain handoff rate for the initially registered VMDs at varying tiers.

fore, User_4 has the mobility reference tier of 4. The remaining 20% of the handoffs are out of domain and handled by the service providers rooted at tier 3. User_4's out-of-domain handoff rate is 0.2.

- *User_3* makes 80% of handoffs within the common anchor clouds rooted at tier 3 in Fig. 5.1. The remaining 20% of the handoffs are handled equally by the mobility agents in the other common anchor clouds, which are rooted at tier 2. For User_3, the mobility reference tier is tier 3, and the out-of-domain handoff rate is 0.2.

The mobile users are assumed to make 100 handoffs that consist in-domain and out-of-domain handoffs. The number of in-domain handoffs is equally divided among the related service providers. The number of out-of-domain handoffs is equally divided among the related service providers. The parameter values stated in Table 5.1 and the network setup in Fig. 5.1 are kept the same.

As it can be understood from the definition of the aforementioned mobile user profiles, the roaming range of the mobile users increases starting from User_6, User_5, User_4, and User_3, in that order. To illustrate, User_6 can be considered a local mobile user roaming within a city, such as Rochester, most of the time; while User_5 is roaming in New York

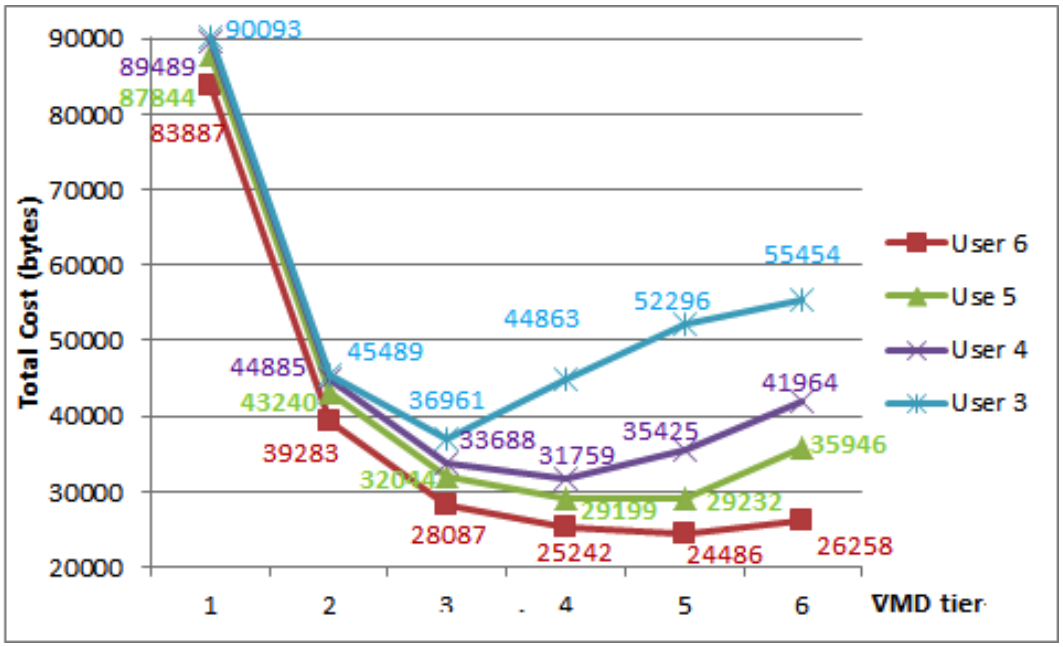


Figure 5.8: Handoff cost vs. initially registered VMD for users.

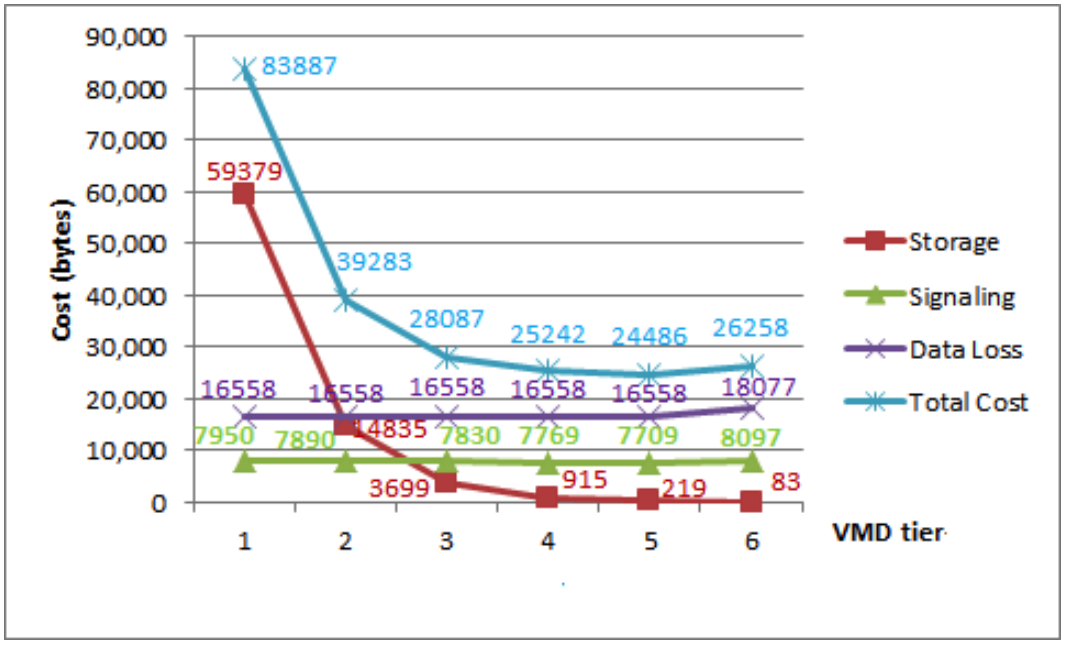


Figure 5.9: Cost vs. the initially registered VMDs for User_6.

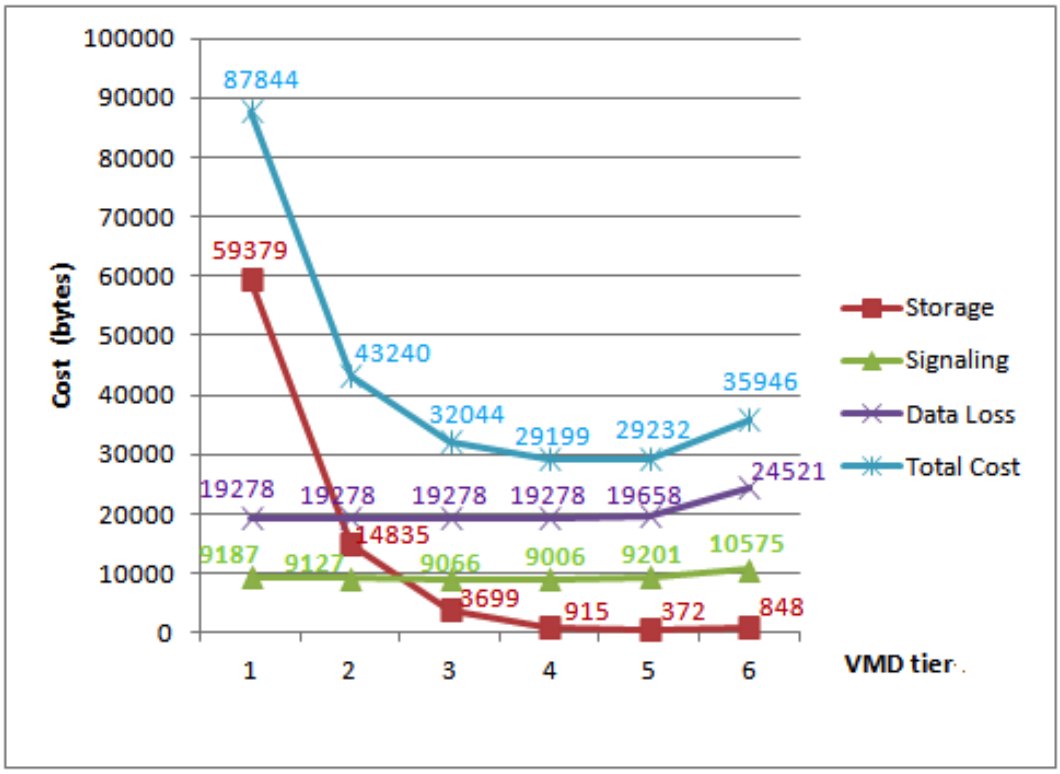


Figure 5.10: Cost vs. the initially registered VMDs for User_5

State; and User_4 is roaming the East cost of the U.S., User_3 can be considered as roaming within an area, such as the United States, where his/her roaming needs are handled by the mobility agents in the common anchor clouds at the upper tiers, compared to the other mobile users. Note that we did not include User_1 or User_2 who can be considered frequent worldwide or cross-continent travelers, respectively. We assume User_1 and User_2 as rare cases. We aim to cover the most typical mobile users to mimic the possible different mobile users with different roaming needs. We want to demonstrate the applicability of the framework. These profiles are not based on any particular mobility model. These mobile user profiles are created to have sample of number of handoffs that are handled by the mobility agents. The number of handoffs are necessary to calculate Eqns. (5.5 and 5.9). The total cost values for User_6, User_5, User_4, and User_3 when they are initially registered to the VMDs at tiers 1 - 6 are illustrated in Fig.5.8. Detailed drawing of User_6's total cost with the cost components: signaling, storage and data loss costs is in Fig. 5.9 - 5.12. We have created the aforementioned profiles to demonstrate the applicability of the framework. In Section 6, you can find our proposed mobility model.

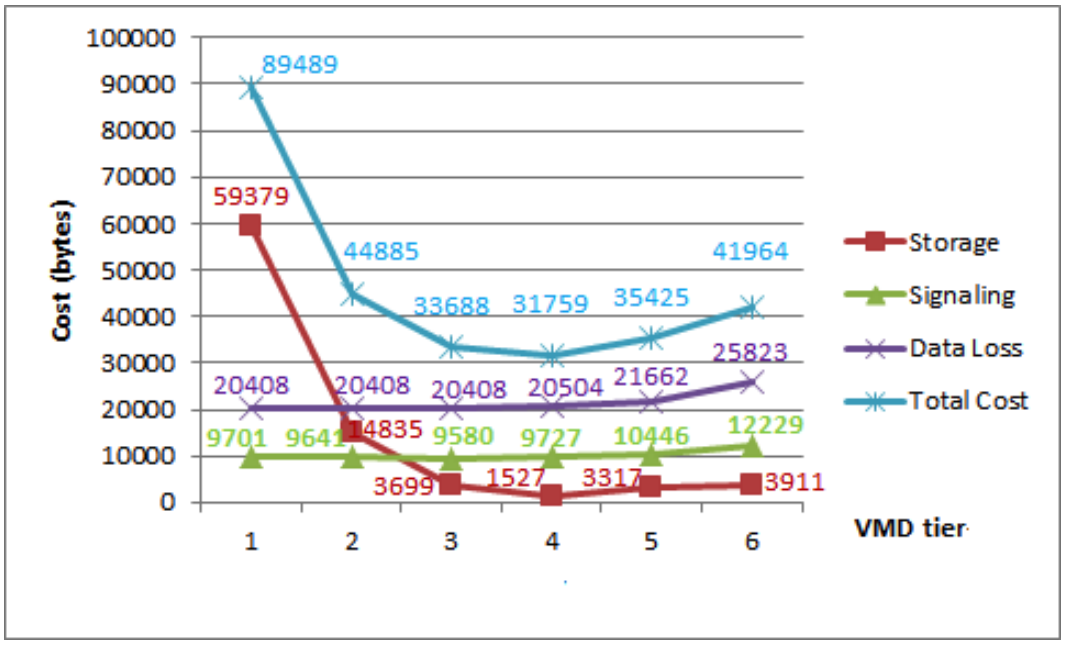


Figure 5.11: Cost vs. the initially registered VMDs for User_4

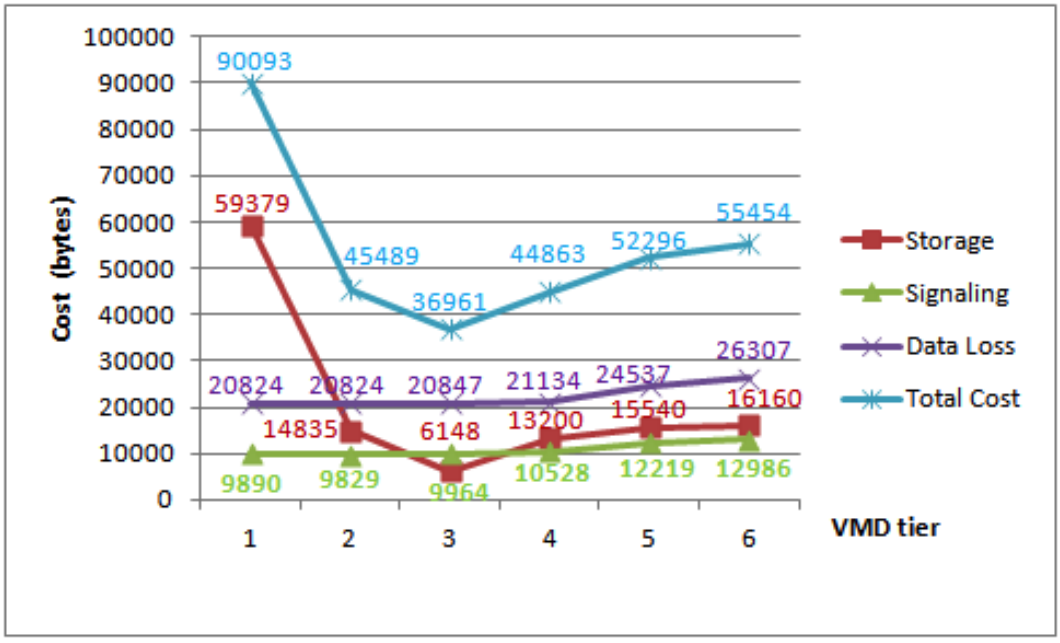


Figure 5.12: Cost vs. the initially registered VMDs for User_3

Fig. 5.9 illustrates the total cost of having a handoff service, calculated by Eqn. (5.1), from the VMDs at tiers changing from tier 1 to tier 6 for the mobile users of which mobility profiles stated above. For User_6, the VMD that gives the minimum total cost (24486 bytes) is the VMD at tier 5 because the minimum signaling cost (7709 bytes) and the data loss cost (16558 bytes) are observed even when the storage cost (219 bytes) is not at its minimum value, as illustrated in Fig. 5.9. The tradeoff for User_6 in registering with the VMD at tier 5, rather than with the VMD at tier 6, is that the mobile user will have 388 bytes less signaling cost, and 1519 bytes less data loss cost, while having 156 bytes more storage cost.

Fig. 5.10 shows the storage, signaling, and handoff costs when User_5 registers initially to VMDs at tier 1 - 6. The optimum VMD that User_5 should register with is the VMD at tier 4 because the minimum total cost (29199 bytes) is incurred in that VMD. As detailed in Fig. 5.10, the optimum VMD tier is one tier above in the topology than the actual mobility reference tier of the mobile user, because the sum of the decrease in the signaling cost and the data loss cost is higher than the increase in the storage cost compared to the case of registering with the VMD at tier 5.

Fig. 5.11 shows the storage, signaling, and handoff costs when User_4 registers initially to VMDs at tier 1 - 6. As illustrated in Fig. 5.11, if the mobile user registers with the VMD at tier 3, rather than the VMD at tier 4, he will have 2172 bytes of extra storage cost while having 147 bytes less signaling cost and 96 bytes less data loss cost. Among the VMDs, the mobile user gets the lowest handoff cost when he registers with the VMD at tier 4. If the mobile user registers with the VMD at tier 4, rather than the VMD at tier 5, he will have 1790 bytes less storage cost, 719 bytes less signaling cost, and 1158 bytes less data loss cost. Compared to the case of being registered with the VMD at tier 5, the mobile user is better off at all of the handoff cost components when he registers with the VMD at tier 4. For User_4, the optimum VMD that he should register with is the VMD at tier 4, which gives the minimum total handoff cost (31759 bytes).

Fig. 5.12 shows the storage, signaling, and handoff costs when User_3 registers initially to VMDs at tier 1 - 6. If f User_3 registers with the VMD at tier 2, rather than the VMD at tier 3, the increase in the storage cost dominates the decrease in the signaling and data loss costs. However, if User_3 registers with the VMD at tier 4 rather than tier 3, he will get less storage, signaling, and data loss costs, as illustrated in Fig. 5.12. The optimum VMD that User_3 should register with is the VMD at tier 3 with the minimum total handoff cost (36961 bytes). User_3's mobility reference tier is the same as the tier of the optimum VMD, as illustrated in Fig. 5.12.

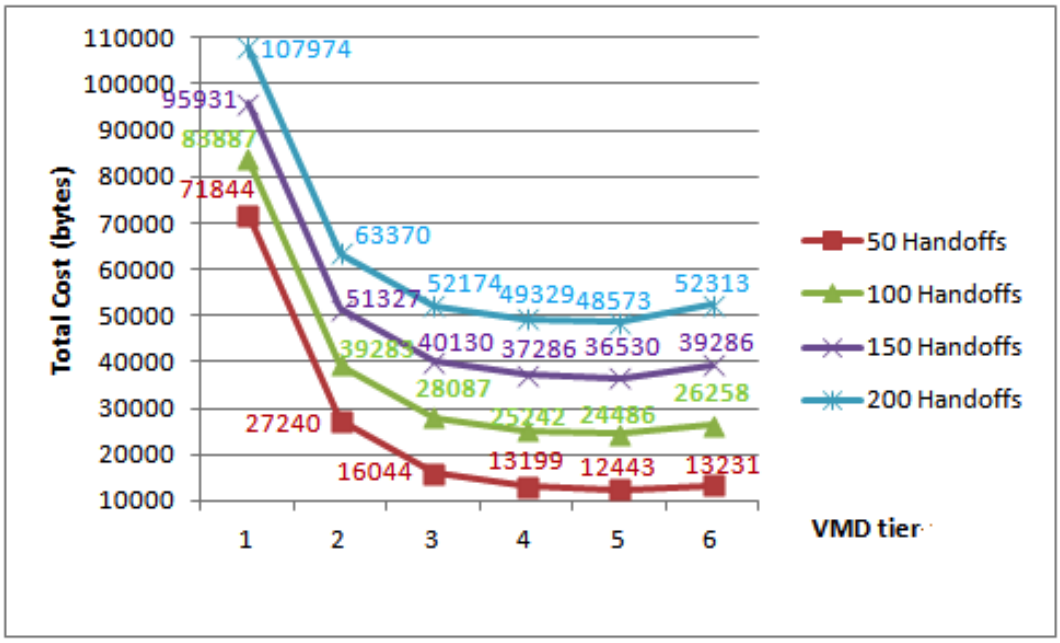


Figure 5.13: Cost vs. the initially registered VMDs for varying number of handoffs of User_6.

5.6.3 Effect of Number of Handoff

In this section, we change a mobile user's number of handoffs and observe the total handoff cost. We expect to see that the mobile user's data loss cost will increase with the increase in the number of handoffs as expressed in Eqns. (5.10 and 5.11). We maintain the parameter values stated in Table 5.1 and the same network setup in Fig. 5.1. We use the User_6 profile as an example. We then calculate the total handoff cost values using Eqn. (5.1) for User_6, whose user profile was explained previously, making 50, 100, 150, and 200 handoffs. Fig. 5.13 illustrates that the total handoff cost values increases with the increasing number of handoffs as it affects the in- and out-of-domain signaling cost and data loss cost linearly as expressed in equations 5.6, 5.7, 5.10 and 5.11. The storage cost is not affected by the number of handoffs, which is in parallel with Eqn. (5.4). The VMD that the mobile user should register with, however, does not change, because the characteristics of the mobility does not change especially within the area that the mobile user roams; the only change occurs in the number of handoffs.

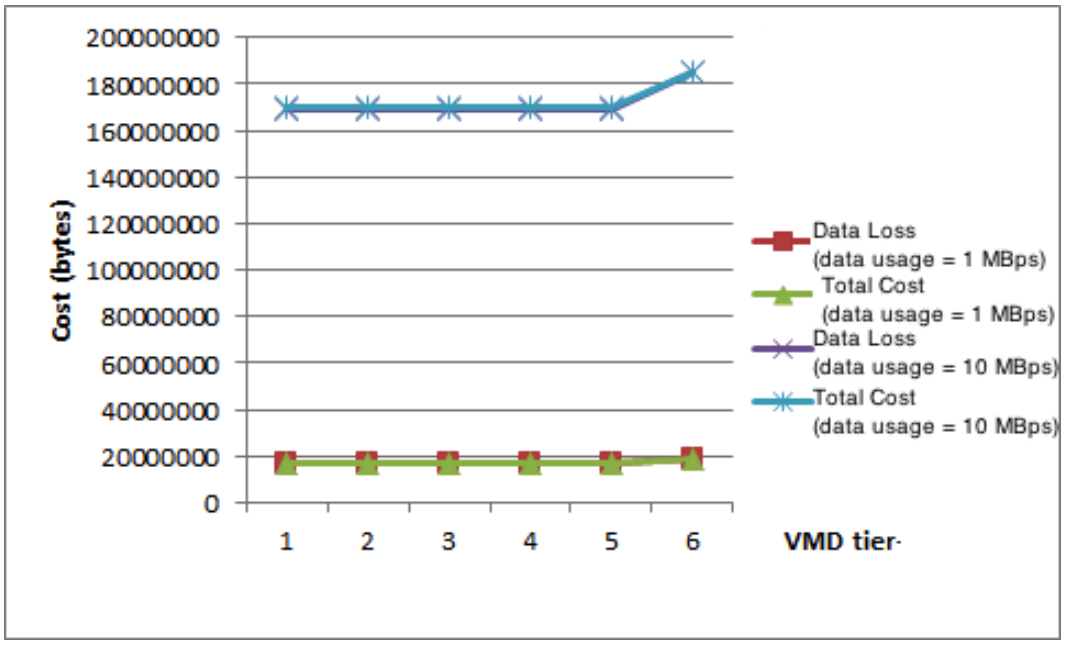


Figure 5.14: Cost vs. initially registered VMD tiers for User_6's data usage 1 MBps and 10 MBps.

5.6.4 Effect of Data Usage

Fig. 5.14 shows the costs for the case (1) where the mobile user's data usage is 1 MBps, which represents the low data usage, and the case (2) where the mobile user's data usage is 10 MBps, which represents high data usage. Data usage on a mobile device changes depending on the number of applications that are running and the amount of data that each of these applications is acquiring from the Internet. The communication of the mobile user on the mobile device may be affected due to handoff latency. We aim to analyze the effect of data usage on data loss cost and total handoff cost for a mobile user. The amount of data usage is $\lambda_s \cdot R_{data}$, as expressed in Eqn. (5.10 and 5.11). In our analysis, we pick User_6 as an example and maintain all of the same parameter values in Table 5.1 except for λ and R_{data}. The mobile user has more data loss cost when the applications he is running need to acquire more data from the Internet. The increase in the data loss cost reflects on the total handoff cost. The trends on the data loss cost and the total handoff cost do not change with the data usage because the number of handoffs and the tier they are handled on ($HO(MA)$ in equations 5.10 and 5.11) is another factor affecting the data loss cost, and they depend on the mobile user's mobility profile. Please note that data loss cost (in the order of millions of bytes) is much higher than storage and signaling costs (in the order of thousands of bytes), hence, in the Fig. 5.14, the total cost and the data

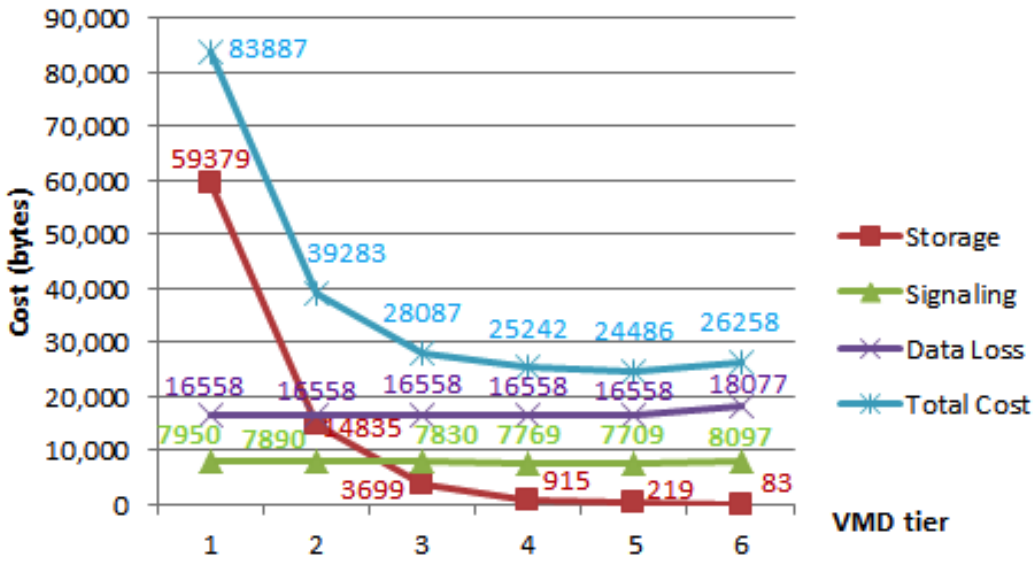


Figure 5.15: Cost vs. initially registered VMDs when $w_p = 0.3$ and $w_d = 0.7$.

loss cost look like they are overlapping. The actual numerical values are not placed in the figure to avoid clutter in the presentation.

5.6.5 Effect of User Sensitivity to Cost Components

Mobile users can have various sensitivity levels to data loss, which is denoted by w_d in the handoff cost framework in Eqn. (5.1). The value of w_d increases with the increasing level of sensitivity to data loss. On the other hand, a mobile user may be sensitive to the extra costs imposed by the service provider, such as for storage and signaling costs, due to the handoffs. The mobile user's sensitivity is denoted by w_p. If the mobile user does not care about the costs imposed by the service provider, we expect the w_p to get lower values compared to the case where the mobile user gives greater importance to the storage and signaling costs imposed by the service provider. We aim to show the significant impact of the sensitivity parameters w_p and w_d on the signaling, storage, and data loss costs. Therefore, we create two cases:

- *Case (1):* $w_p = 0.3$ and $w_d = 0.7$, which represents a case where the mobile user gives more importance to his/her data communication rather than to the extra costs that may be imposed by the service provider due to the storage usage or signaling overhead.
- *Case (2):* $w_p = 0.7$ and $w_d = 0.3$, which represents a case where the mobile user gives

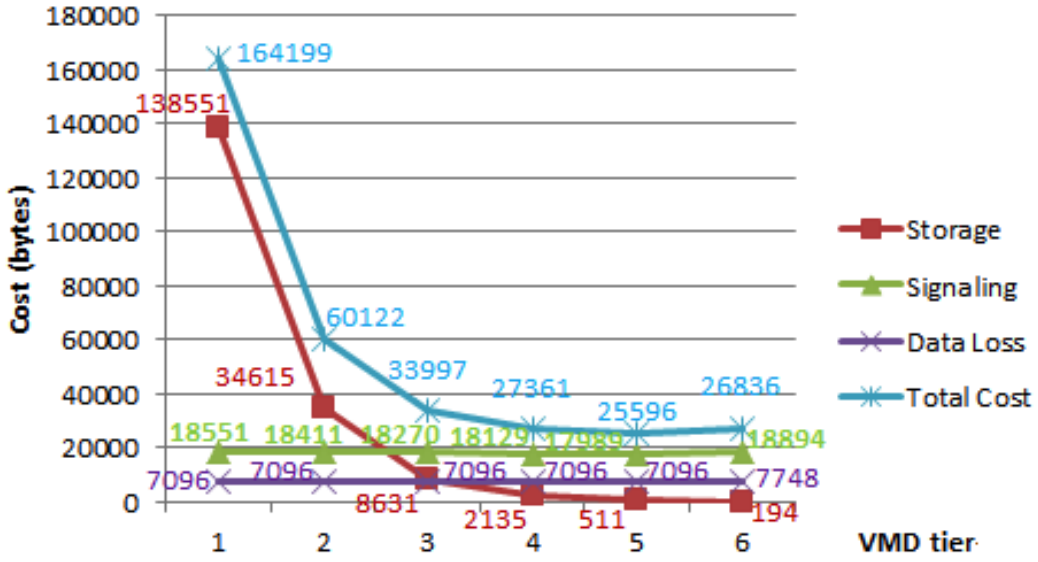


Figure 5.16: Cost vs. initially registered VMDs when $w_p = 0.7$ and $w_d = 0.3$.

more importance to the costs imposed by the service provider compared to the data loss.

To illustrate, we pick User_6 as the subject. He makes 100 handoffs, 80% of them are within AS 6.1, and the remaining handoffs are handled in equal amounts by mobility agents in ISP 5.1 and its customer ASes in Fig. 5.1. The parameter values in Table 5.1, except w_p and w_d, are maintained. Fig. 5.15 and 5.16 illustrate that changing w_p and w_d values does not affect the patterns on the storage cost, signaling cost, and data loss cost curves because these costs are mainly dependent upon the tier of the initially registered VMD, the tier of the common anchor cloud, and the handoff amount that the mobility agent in the common anchor cloud handles, as expressed in Sections 5.2.1, 5.2.2 and 5.2.3. Since the mobility profile of User_6 does not change, the optimum VMD also does not change. However, the values of the cost components are affected due to the change in the sensitivity weights, w_p and w_d. Comparing case (2) to case (1), the signaling cost and the storage cost increase while the data loss cost decreases. Therefore, the total cost is affected more by the storage and signaling costs. To illustrate, the curve that connects storage cost values in Fig. 5.16 is steeper than the one in Fig. 5.15, and the same trend is observed in the total cost curves in both of the figures. In the case of higher w_p, compared to w_d, the storage and signaling costs dominate the total handoff cost value. The total cost is different for each initially registered VMDs. This amount of total cost difference changes

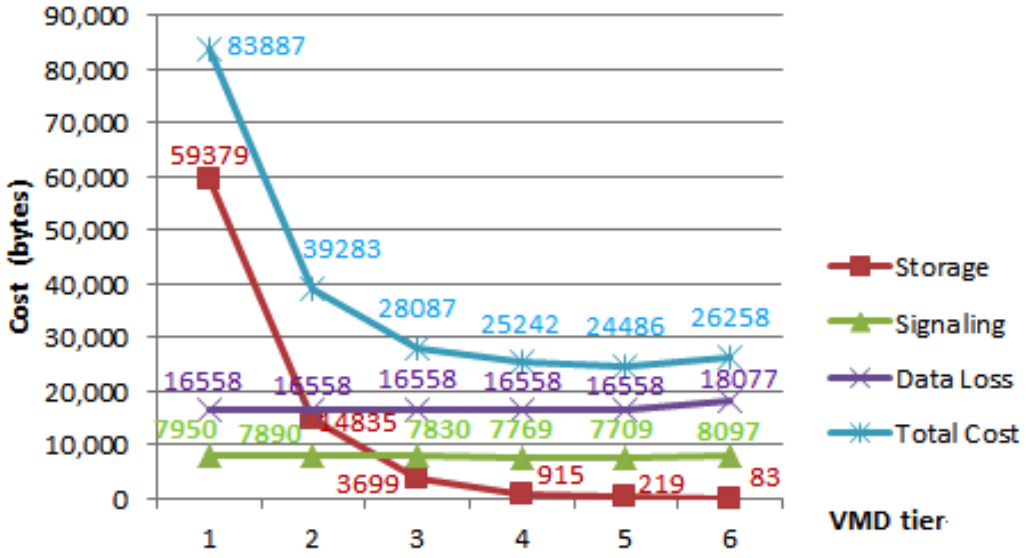


Figure 5.17: Cost vs. initially registered VMDs when $\tau = 1.1$.

from case (1) to case (2), as seen in Fig. 5.15 and Fig. 5.16.

In Fig. 5.15 and Fig. 5.16, the storage cost is high for the cases where the mobile user is registered with VMDs at tier 1. One of the reasons for the high storage cost is that when a user registers with a VMD at tier 1, the mobile user profile information is stored in anchor points in the topology as expressed in Eqn. (5.12). The storage cost decreases significantly when a mobile user initially registers with VMDs at higher tier values in the topology because the number of the mobile agents inside the VMD decreases. Further, the high w_p value causes the high storage cost, too.

5.6.6 Effect of External Service Cost Multiplier

A mobile user may cross the boundaries of the VMD with which he is initially registered, as seen in the previous sections. In that case, the mobile user needs to register temporarily to a different VMD. The temporarily registered VMD will store the user-related data in its proxy AAA servers and forwarding bases to be able to give handoff service, as expressed in Eqn. (5.4). During the handoff of the mobile user within the new VMD, there will be extra signaling overhead on the new service provider as expressed in Eqn. (5.5). The storage and signaling costs that are incurred on the new VMD, which will be imposed on the mobile user, may not be similar to the costs in the initially registered VMD because the mobile user is temporarily registered with the new VMD. The extra cost of getting

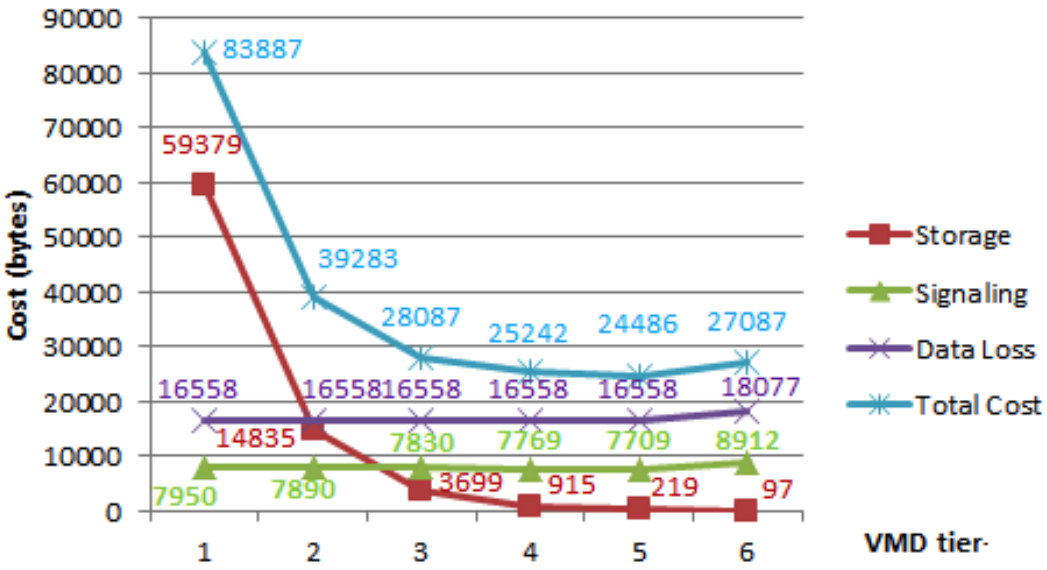


Figure 5.18: Cost vs. initially registered VMDs when $\tau = 1.5$.

services temporarily from a VMD is represented by the cost multiplier τ as expressed in Sections 5.2.1, 5.2.2 and 5.2.3. In this subsection, we aim to analyze the effect of τ on the cost components and the VMD under which the mobile user should register.

We discuss the costs User_6 has for two cases: when (1) $\tau = 1.1$ and (2) $\tau = 1.5$. The reason for picking values higher than 1 is that we assume that a temporarily registered service provider is more costly compared to the one under which the mobile user is registered permanently. Depending on the business relationship between the service providers, the maximum value of τ can vary. In case (1) and case (2), we do not place the τ values so high that they will dominate the cost of the service retrieved from the permanently registered service provider. The other parameter values in Table 5.1 are maintained. The numerical results are presented in Fig. 5.17 and Fig. 5.18.

Fig. 5.17 and Fig. 5.18 illustrate that the change on τ does not affect the cost components for the cases where User_6 is initially registered with the VMDs at tiers 1 to 5 and in the cases where the mobile user does not make any out-of-domain handoffs and, hence, no temporary storage or signaling costs are incurred. However, the storage and signaling costs at VMD 6 are higher in Fig. 5.18 compared to the one in Fig. 5.17 because of the higher cost of getting service from the temporarily registered VMD. The increase in the signaling and storage costs also reflects on the total handoff cost and makes being registered with the VMD at tier 6 more costly. If User_6 registers initially to the VMD at

tier 5, rather than the VMD at tier 6, the mobile user pays for 388 bytes of signaling but gains 136 bytes of storage, as presented in Fig. 5.17, while the mobile user is better off paying for 1203 bytes in signaling and gains only 122 bytes of storage cost, as presented in Fig. 5.18. These results signify that for User_6, registering with the VMD at tier 6 is more costly, in the case of $\tau = 1.5$, which means that the out-of-domain handoffs are more costly. The optimum VMD that the mobile user should register at is the VMD at tier 5, regardless of $\tau = 1.5$ and $\tau = 1.1$, because the user's mobility profile does not change. He continues to make 80% of the handoffs in AS 6.1 and 20% of the handoffs within AS 6.2, AS 6.3, AS 6.4, and ISP 5.1 in Fig 5.1.

5.7 Summary

We present a novel, user-centric handoff cost framework that includes (i) the network storage cost of user-profile data for providing handoff service, (ii) extra signaling overhead that is incurred in the system during the user's mobility support, and (iii) the user's loss of data communication because of handoff latency. This framework enables us to observe the characteristics of the dynamics affecting the handoff cost from a user's perspective. Leveraging the user-centric mobility support given by the VMD, this framework is applied to help choose the mobility domain that is less costly for a user, depending on his/her mobility preferences and requirements. We conduct a numerical analysis for different user profiles, different data usage capacities, cost sensitivities, and service-provider related parameters introduced in the framework. We identify the optimal VMD for a user with given preferences and parameters, and we then discuss why that VMD is optimal domain. We also demonstrate that the handoff cost framework can be applied to IPv6-based protocols to show its applicability to other protocols besides VMD-based protocol.

Chapter 6

Optimization of Handoff Cost

Mobile users have different mobility characteristics. For instance, mobile users roam with varying speeds, for different time durations, and have diverse roaming ranges. During roaming, a mobile user may need to leave the current access point and connect to a new access point. In the meantime, the mobile user's ongoing communication over the Internet needs to be transferred to a new access point. De-registration of the mobile user from the current access point in order to register to the new access point, and the transfer of ongoing communication of the mobile user by route updates and state changes on the network devices is called handoff [143]. Handoff-related operations require signaling, messaging, storage of user-related data, and packet delivery. The handoff process also consumes network resources; therefore, the activities impose additional costs to both the service provider and the mobile user.

Mobile users transfer various amounts of data to and from the Internet with their mobile computing devices. A possible delay during handoff might cause data loss in communication, which may be unacceptable to the mobile user. Depending on the roaming trajectory and the speed, the mobile user may make several handoffs - which can be highly disruptive to the mobile user's communications. For the least disruption and the best quality of service, a mobile user has to obtain connectivity services from service providers that fit the mobile user's roaming habits and have the optimum handoff cost formulated in Eqn. (5.1).

Service providers provide resources for different services, which include registration and storage of the mobile user's profile and credentials in proxy AAA (authentication, authorization, and accounting) servers (See Section 5.1.1 for details.). Further, each handoff requires signaling on the system to update routing entries and to deliver the data packets to the mobile user at the current point of connection (See Section 5.1.2 for details.). Service providers may prefer to use minimum resources to offer an acceptable roaming service to

a mobile user.

Today, mobility service providers offer service packages that are priced/marketed mostly on call duration, texting, and data usage. The information regarding how these packages are created, what the cost components are, and what affects the price of these packages are not readily available to mobile users. The services do not provide mobile user-centric customization. The optimization studies in the literature (See Section 2.6) are also service-provider centric rather than user-centric. The service-provider centric approach is reasonable because the service providers are the ones that invest on the infrastructure. There are also architectural constraints that prevent a user-centric approach. The current Internet architecture does not provide flexibility for users to select any mobility domain because and they are managed by different protocols such as MIPv6, HMIPv6, and PMIPv6. Users need to register to each domain separately (See Section 2.2 for details). So it is natural for researchers not to pursue optimization from the user point of view.

We envision a mobility ecosystem that allows a mobile user to select the optimum mobility domain that aligns with the mobile user's roaming needs and quality-of-service requirements with minimum handoff costs. A user-centric, Virtual Mobility Domain (VMD) architecture provides such a system by providing overlapping domains with different size of coverage areas managed by the same protocol. By registering to an optimum domain, the user will get signaling, storage, and data loss cost that aligns with his cost sensitivities. See Eqn. (6.17) for cost formulation. The crucial point to grasp here is that the proposed architecture is fixed and the user 's optimum choice will not incur further costs on the service provider. Therefore, we believe mobile users satisfaction should be one of the goals of the service providers because mobile users are the end-consumers of the mobility services. We expect customer-oriented services get higher attention, ultimately increasing the profit of the service provider. User-centric approach is also in favor of the service provider because the optimum domain results in less handoff cost consisting of storage and signaling costs due to consumption of network resources. See Eqn. (6.17). We also aim to differentiate our work from the existing literature by studying a non-explored area based on employing VMD.

A powerful feature of VMD is its capability to provide user-centric mobility management. VMD introduces a *virtual network cloud* concept where a mobile user gets an address from a virtual network cloud that is not constrained to any physical network cloud. A physical network cloud can be an Internet service provider (ISP) network; a point of presence (POP) in an ISP; an autonomous system (AS); or a set of backbone, distribution, or access routers in a network. A virtual network cloud defines the boundary of a VMD. To illustrate, we provide a few VMD scenarios in Fig. 6.1. We deploy VMD 3 under ISP C. VMD 3 is suitable for mobile users who mostly roam within ISP C, and this covers AS

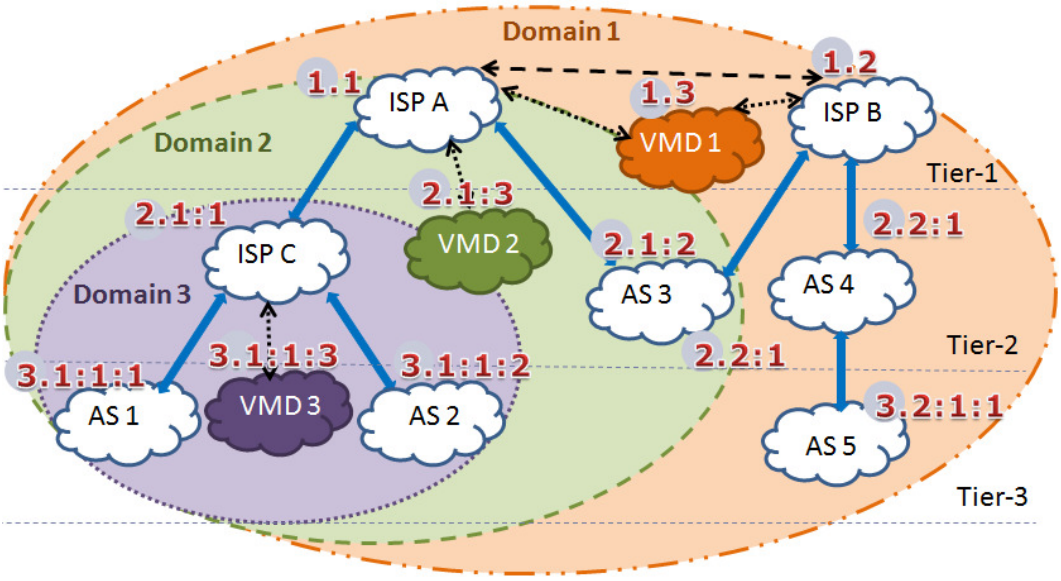


Figure 6.1: The multiple overlapping Virtual Mobility Domains are applied on the Floating Cloud Tiered Internet working model. Dotted arrows show the VMDs' deployment to upper-tier clouds.

1 and AS 2. In Fig. 6.1, VMD 3 is shown as domain 3 within the dotted circle. When we deploy a VMD to the upper tiers, such as VMD 2 or VMD 1, the scope of the mobility domain expands to domains 2 or 1, respectively. The overlapping mobility domains in the proposed mobility architecture offer mobile users the flexibility of registering under any mobility domain depending upon their mobility patterns. If a mobile user roams frequently within a wider area, then we recommend a VMD at an upper tier. The registration and storage costs, signaling overhead, and packet delivery cost may change depending on the tier of VMD with which the mobile user registers. Once a mobile user registers to a mobility domain, the mobile user will be assigned an address that the mobile user can use for roaming across all physical network clouds that are covered under that domain. Therefore, the mobile user will not have to get a new address if the mobile user changes point of connection. In this way, repetitive registration costs, signaling overhead, and handoff delays can be reduced.

To illustrate, VMD 2 has an address 2.1:3, because it is deployed under ISP A, which resides in tier 1. A mobile user registered under this VMD cloud will thus get an address in tier 3, namely 3.1:3:n (where n can be a unique integer value). Mobility agents in the common anchor clouds (CACs) support a mobile user's mobility in VMD. For instance, if a mobile user moves from AS 2 to AS 3, then ISP A will be the common anchor cloud,

as it is the lowest common point in the hierarchy between the ASes in the architecture. A mobility agent in ISP A will handle the mobile user's handoff by communicating with the necessary physical network clouds. As we confirm with performance studies in Chapter 4, this collaborative network-based, handoff management approach imposes fewer costs on mobile users with less signaling overhead and handoff latency.

We present optimization studies in the current literature prior to our approach. The handoff cost that is incurred in HMIPv6 and PMIPv6, with all the sub processes that cause signaling overhead and handoff delays, are extensively studied [39, 42, 110]. Other researchers aimed to decrease network resource usage by optimizing the processes involved in location tracking, mobility control messaging, and packet delivery [104, 113, 114]. They proposed new algorithms, network elements, and mobility agents to track the location of a mobile user with less network resources and power consumption [104, 113, 114]. In the proposed VMD architecture, we demonstrate that the collaborative network-based, mobility-management scheme performs better in mobile user tracking and packet delivery in comparison with HMIPv6 and PMIPv6 for the cases presented in Chapter 4. In this work, we conducted an optimization study on the proposed mobility architecture by finding the optimum domain that a mobile user needs to register with when considering mobile user-related mobility parameters.

Jeon et al. [116] and Dutta et al. [118] aimed to improve the handoff initiation by applying reactive and proactive handoff approaches. Jeon et al. and Dutta et al. further aimed to decrease both packet-delivery-related costs and handoff delays. We do not propose a new handoff initiation algorithm. Our optimization study focuses on finding the best mobility domain among those available in the VMD architecture that are aligned with the mobile user's preferences.

Pack et al. [109] and Vilhar et al. [115] proposed optimum network topology that minimizes location-update and packet-delivery costs. We build the proposed mobility architecture on the tiered structure existing among ISPs and ASs. Users can register with any tier for their mobility support. Our optimization study focuses on finding the optimal tier that a mobile user should register with. The optimal tier is the tier that brings minimum handoff costs and satisfies the mobile user's roaming preferences.

In [119, 121], researchers aimed to provide user-centric mobility support by creating handoff policies that included network-related metrics, application quality-of-service (QoS) requirements, and mobile user preferences. We do not impose handoff policies; instead, our optimization tool accepts the mobile user parameters to allow a mobile user to decide which tier a mobile user should register with based on his/her preferences. See the related research work on the optimization of handoff costs in the current Internet

architecture in Section 2.6.

We aim to find the optimum mobility domain that the mobile user should register with to experience minimum handoff cost in the VMD architecture. There are several handoff cost components such as storage of mobile user-related data, registration cost, signaling overhead, and data loss. In Section 5.1, we proposed a unified handoff cost framework that combines all the handoff cost components under the same metric. As we follow a user-centric approach, the handoff cost framework recognizes a mobile user's roaming characteristics, such as frequency of handoff, mobile data usage, relative sensitivity to each cost component, as well as tolerance to possible data loss. Finding an optimum mobility domain, among several overlapping domains with various sizes, that aligns with the mobile user's mobility preferences may pose a challenge to the mobile user and the service provider. Refer to Section 6.2.1 for how we approach this challenge.

As a preliminary work to this optimization study, we analyzed the number of handoffs that a mobile user makes and the location of the mobility agents in the VMD architecture that handles the mobile user's handoffs. Therefore, we are required to have a mobility model to retrieve this information. In Section 6.1, we study a mobility model that represents a mobile user who visits mostly around the center of his roaming region. We focus on the mobile user's handoff frequency, roaming range, and speed of movement.

Our proposed mobility model differs from the mobility models in the literature. The mobility models in the literature (presented in Section 2.5) did not consider the centered-movement characteristics of the mobile user, which we propose in our model. For example, a fluid-flow mobility model [108] assumes that the mobile user is going out of a region all of the time, which is proportional to the population density in any given area. In our proposed mobility model, a mobile user has a tendency to move in the center of a roaming area. We do not consider the group-based mobility models, either the nomadic-community model or the pursued-mobility model [101], because we assume a mobile user makes movement decisions independently from a group's motion. Our mobility model differs from the random walk mobility models [102] in terms of homogenous distribution of movement directions and the consideration of the average speed of the mobile user. We build our proposed mobility architecture on the already-existing tiered structure between ISPs and ASs in the Internet, which are not limited to residential or business buildings. Therefore, in our mobility model, we do not consider the effect of the buildings or roads in contrast to the map-based mobility models [106,107].

In Section 6.2.1, we model the handoff cost framework that was presented in Section 5.1 as an optimization problem so that the Eqn. (5.1) becomes the objective function of our optimization problem. We discuss the characteristics of the optimization problem

by checking if the objective function is linear or nonlinear, and if the decision variable is integer or non-integer. This analysis helps us to choose the optimization tool in the literature to efficiently solve the handoff cost optimization problem that we formulated. In Section 6.3.2, we use the optimization software GAMS [144] to solve our optimization problem. By solving the optimization problem, we can find the optimum domain for any type of mobile user with any mobility characteristics in any network settings.

The contributions of this chapter are the following:

- To the best of our knowledge, our mobility study in Section 6.1 is the first of its kind at providing a guide for deriving the number of handoffs in a typical VMD and identifying the common anchor clouds where the mobility agents that handle the mobility reside. We propose a mobility model that is also the first of its kind that represents the mobile user's centered-movement characteristics.
- We solve a handoff cost optimization problem considering a mobile user as the primary focus in Section 6.2. We identify the type of the handoff cost optimization problem (for example, whether it is nonlinearity or linear). We then decide on suitable optimization algorithm to solve the problem.
- We conduct numerical validation of the mobility and optimization studies in Section 6.3. We model the handoff cost problem using optimization tool. We conduct a numerical study to find out the optimum VMD for a given user-mobility model.

6.1 A Study on Finding Number of Handoffs

The number of a mobile user's handoffs is a very important factor in handoff cost optimization because it affects the handoff cost and the optimum VMD that a mobile user registers to, as stated in the introduction section. $HO(MA)$ expressed in Eqns. (5.6), (5.7), (5.10) and (5.11) denotes the number of handoffs handled by the mobility agent in the common anchor cloud. In this section, we aim to provide a method by which $HO(MA)$ can be obtained from a given mobility model.

A mobility model includes the distribution of the locations that a mobile user visits and the number of handoffs. The number of handoffs is calculated as multiplication of handoff rate and time that the mobile users movement is observed. A mobility function is an element of a mobility model because it characterizes the probability distribution of the locations that the mobile user visits. In the following paragraph, we will propose a mobility model and present the mobility characteristics that we want to analyze. We will focus on the proposed mobility function. The aim is to get the probability distribution of the locations that the mobile user visits. We will also identify the common anchor

clouds while the mobile user visits those locations. Later, including other elements of the mobile user's mobility model such as handoff rate and movement duration, we will find out the number of handoffs that are handled by mobility agents in common anchor clouds.

We propose a mobility model that represents a mobile user who spends most of the time in the center of the roaming region and spends less time on the edges of the roaming area. Our intuition is that a person moves most frequently to specific locations (such as home, school, work, and recreational areas), and rarely goes to different locations involving great distances (such as overseas for a vacation). In our proposed mobility model, we assume that the frequently visited places are close to the center of the mobile user's roaming area, while the rarely visited places are close to the edges of the roaming area. Note that we are not interested in the individual steps that the mobile user takes while moving in the roaming area. Instead, we are interested in the probability of finding the mobile user at a specific part of the roaming area after a steady state is reached (i.e., after a long time the mobile user has started to move). We also consider the mobile users who have different handoff rates.

As part of our mobility model, we propose a radial mobility function $m_a(r)$:

$$m_a(r) = \begin{cases} \left(1 - \frac{r^2}{2a^2} + \frac{r^4}{4a^4} - \frac{r^6}{6a^6}\right) \cdot \left(1 - \frac{r^2}{a^2}\right)^2 & \text{for } 0 \le r \le a \\ 0 & \text{elsewhere,} \end{cases} \tag{6.1}$$

where a is corresponding to roaming radius. For example, $m_1(r)$, $m_2(r)$, and $m_4(r)$ denote radial mobility functions for mobile users having roaming radius values of 1, 2, and 4, respectively.[1] These functions are shown in Fig. 6.2.

We now define a two-dimensional mobility function $F_a(x, y)$, in the Cartesian coordinate plane, as follows:

$$F_a(x, y) = m_a(r) \quad \text{where} \quad r = \sqrt{x^2 + y^2} \tag{6.2}$$

The Cartesian mobility function $F_a(x, y)$ describes a mobile user's roaming region, in this case, a circular region of radius a where the position $(0, 0)$ is taken to be the mobile user's initial "home" position.[2]

[1] We wrote the mobility function in polar coordinates to emphasize the circular symmetric nature of the mobility function meaning dependency on the distance from the origin of the roaming area but independency from the angle.

[2] We formulated the mobility function in Cartesian coordinates to easily illustrate how the movement of the mobile user results in handoff in the FCT internetworking model that is mapped to a Cartesian coordinate plane in Section 6.1.1.

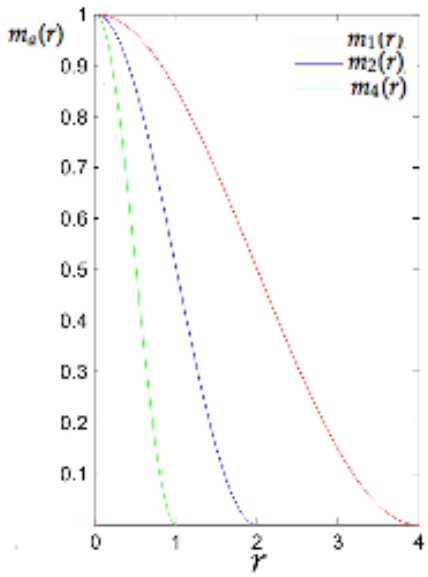


Figure 6.2: The mobility function $m_a(r)$.

Let V_a denote the volume under the graph of $z = F_a(x, y)$. After converting the double integral from Cartesian to polar coordinates, the double integral can be expressed as the product of two single integrals, as follows:

$$V_a = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_a(x,y)\mathrm{d}x\mathrm{d}y = \int_0^{2\pi}\int_0^a m_a(r)r\mathrm{d}r\mathrm{d}\theta = \int_0^{2\pi}\mathrm{d}\theta \cdot \int_0^a m_a(r)r\mathrm{d}r \tag{6.3}$$

Next, we write:

$$V_a = 2\pi \cdot A_a \text{ where } 2\pi = \int_0^{2\pi}\mathrm{d}\theta \text{ and } A_a = \int_0^a m_a(r)r\,\mathrm{d}r. \tag{6.4}$$

Now, for any given radius value a, we can construct a probability density function in the Cartesian plane as follows:

$$f_{XY}(x,y) = \frac{1}{V_a}F_a(x,y) = \frac{1}{2\pi A_a}F_a(x,y) \tag{6.5}$$

Specifically, this means that the probability of a mobility location $P(x, y)$ being in the infinitesimal rectangle $[x, x + \mathrm{d}x]x[y, y + \mathrm{d}y]$ is given by:

$$P(x \le X \le x + \mathrm{d}x, y \le Y \le y + \mathrm{d}y) = \frac{1}{2\pi \cdot A_a}F_a(x,y)\mathrm{d}x\mathrm{d}y. \tag{6.6}$$

When converting to polar coordinates, which is convenient because of radial symmetry, we have:

$$\frac{1}{2\pi \cdot A_a} F_a(x, y)\mathrm{d}x\mathrm{d}y \rightarrow \frac{1}{2\pi \cdot A_a} m_a(r) r \mathrm{d}r\mathrm{d}\theta \tag{6.7}$$

Because the coordinates are now r and θ, this expression becomes the probability of a mobility location $P(r, \theta)$ being in the infinitesimal rectangle $[r, r + \mathrm{d}r]x[\theta, \theta + \mathrm{d}\theta]$. Thus,

$$P(r \leq R \leq r + \mathrm{d}r, \theta \leq \Theta \leq \theta + \mathrm{d}\theta) = \frac{1}{2\pi \cdot A_a} m_a(r) r \mathrm{d}r\mathrm{d}\theta \tag{6.8}$$

In polar coordinates, unlike for Cartesian coordinates, this two-dimensional probability density function is *separable*, meaning that it can be expressed as a function of r times a function of θ. The useful way to do this is:

$$\frac{1}{2\pi \cdot A_a} m_a(r) r \mathrm{d}r\mathrm{d}\theta = \left(\frac{1}{A_a} m_a(r) r \mathrm{d}r\right) \cdot \left(\frac{1}{2\pi}\mathrm{d}\theta\right) \tag{6.9}$$

So, suppose we let:

$$f_R(r) = \frac{1}{A_a} m_a(r) r \quad \text{and} \quad h(\theta) = \frac{1}{2\pi} \tag{6.10}$$

Then, $f_R(r)$ is the probability density function for the random variable r, meaning that the probability of a random radius falling in the interval $[r, r + \mathrm{d}r]$ is equal to $f_R(r)\mathrm{d}r$ for $0 \leq r \leq a$.

Likewise, $h(\theta)$ is the probability density function for the random variable θ, meaning that the probability of a random angle falling in the interval $[\theta, \theta + \mathrm{d}\theta]$ is equal to $h(\theta)d\theta$ for $0 \leq \theta \leq 2\pi$. The fact that $h(\theta)$ is a constant simply means that all values of θ are equally likely, which is a uniform distribution. [3]

In later sections, the two-dimensional Cartesian probability density function $f_{XY}(x, y)$ will be used to compute segment integrals for determining handoff costs. In Section 6.1.4, the one-dimensional radial probability density function $f_R(r)$ will be used in creating random points that the mobile user visits in a Monte Carlo simulations.

6.1.1 Mapping the FCT Internetworking Model

We provide a mapping algorithm, where we map the FCT internetworking model on a Cartesian coordinate plane to determine the number of handoffs.[4] We also use the

[3] Check Appendix B for the validation of $f_R(r)$ and $h(\theta)$.

[4] Note that the VMD architecture is deployed on the FCT internetworking model, hence, we study the mapping of the FCT internetworking model.

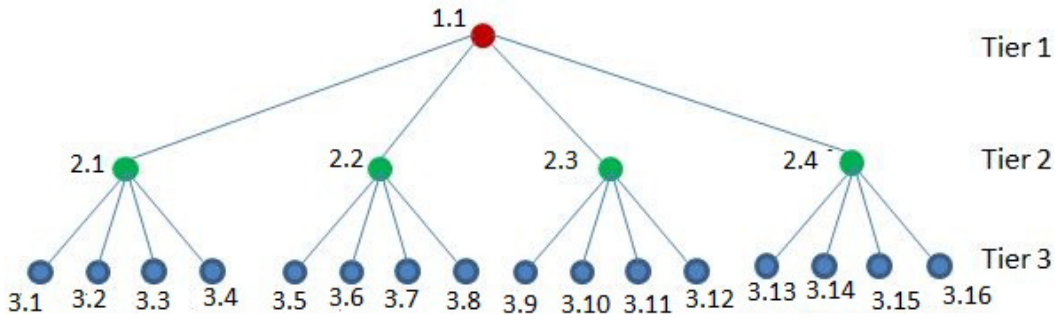


Figure 6.3: The FCT internetworking model.

mapping algorithm to predict where the common anchor clouds, where the mobility agents handling the handoffs, reside. The mapping algorithm is general, hence it can be applied to any tree regardless of the number of children and height. We apply the FCT internetworking model, shown in Fig. 6.3, on a balanced tree. In this balanced tree, each node has γ children and the height is two tiers. Each node in the tree topology represents a cloud in the FCT internetworking model that is an AS or an ISP [139]. We map this balanced tree on a square area with a side length of $2b$, as shown in Fig. 6.4.[5] We set the square on the coordinate plane where the center of the square is (0,0). The clouds of the FCT internetworking model are denoted with $A.B$, where A denotes the tier value that the cloud resides within the architecture, and B denotes an identifier that differentiates the cloud among the peer clouds. We place the root cloud of the FCT internetworking model, which is cloud 1.1, at the center of the square. The children of cloud 1.1, which are clouds 2.1, 2.2, 2.3, and 2.4, are the tier-2 clouds. The tier-2 clouds are located at the center of the quadrants as seen in Fig 6.4. Horizontal (as well as vertical) Euclidian distance between the child cloud and the parent cloud is calculated as follows:

$$d(\psi) = \frac{2 \cdot b}{2^{\psi}} \text{ for } \psi > 1 \text{ and } b \in \mathbb{Z}^{+}, \tag{6.11}$$

where ψ denotes the tier of the child cloud. For example, the Cartesian coordinate of cloud 2.1 in Fig. 6.4 is $(-d(2), d(2))$.

The square area that the FCT internetworking model is mapped onto consists of quadrants as seen in Fig 6.4. We use the same color for the cloud and the boundaries of the quadrants that are denoted by fine strips. For example, cloud 1.1 is in red, and the boundaries of the neighbor quadrants are drawn with red solid strips.[6] The mobility agent in

[5] In the scope of this study, we assume the mobile user does not go beyond this area.

[6] Starting from Section 6.1.2, we will refer to the red strips as strips that belong to cloud 1.1 for simpler explanation and readability.

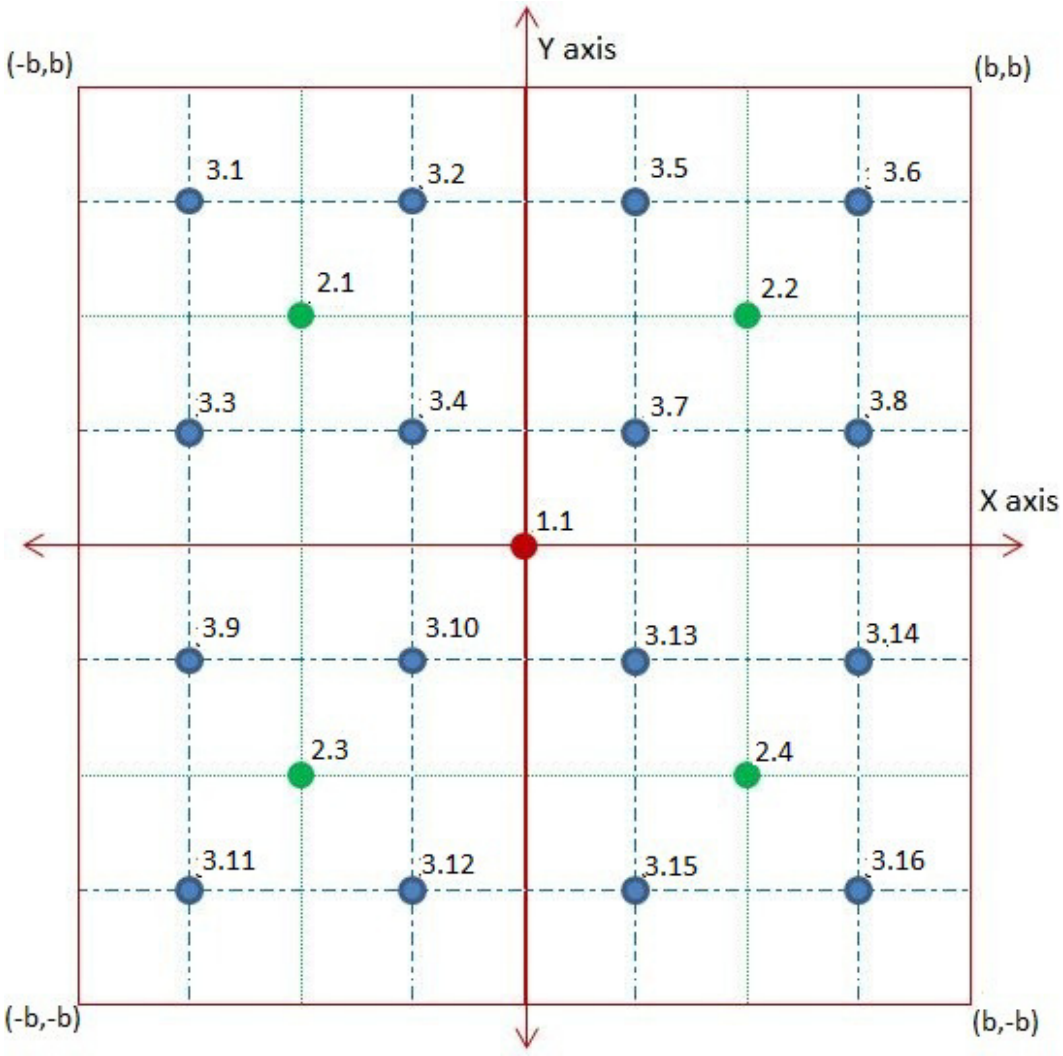


Figure 6.4: The FCT internetworking model is mapped on a Cartesian coordinate plane.

cloud 1.1 handles the handoffs that occur as a result of a mobile user's movement between the quadrants with solid red strips. Cloud 1.1 becomes the common anchor cloud between the previous and the next location of the mobile user. We locate each tier-2 cloud in the center of a square that has side lengths of b, which is composed of four smaller squares. The boundaries of these squares are drawn with green dotted strips and the tier-2 cloud is also green. A mobility agent in a tier-2 cloud is responsible for a mobile user's handoff due to crossing of neighboring squares with green dotted strips. The same mapping and coloring concept is valid for all of the clouds belonging to other tiers.

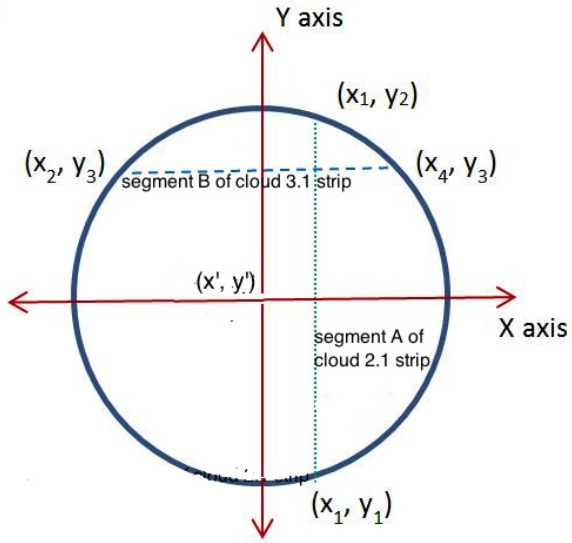


Figure 6.5: The roaming region of a mobile user is a circle where segments A and B represent the strips belonging to clouds 2.1 and 3.1, respectively.

6.1.2 Finding the Probability of Crossing a Boundary

In this subsection, we will formulate the probability of the mobile user coming into contact with (or being on) a boundary strip while moving in the roaming region. We assume that a mobile user makes a handoff when he is exposed to a boundary strip. For illustration purposes, we will use Fig. 6.5, which demonstrates a circular roaming region with origin (x', y'). The origin of the mobile user's roaming region can be located at any part of the plane that the FCT internetworking model is mapped to in Fig. 6.4.[7] Segment A and B in Fig. 6.5 symbolize the strips of clouds in Fig. 6.4. Let segment A be $x = x_1$, $y_1 \leq y \leq y_2$

[7]The origin of the mobile user's roaming region should not be confused with the origin of the area that the FCT internetworking model that is (0,0).

where (x_1, y_1) and (x_1, y_2) lie on the circle.[8] Let's assume that segment A is the only segment that belongs to the cloud 2.1. Let $P(A)$ denote the probability of the mobile user crossing segment A. We formulate $P(A)$ as follows:

$$P(A) = \int_{y_1}^{y_2} \frac{1}{V_a} F_a(x_1, y) \mathrm{d}y \, \Delta x \, , \tag{6.12}$$

where Δx indicates a small positive thickness for the fine strip associated with segment A. The number of handoffs that the mobile user makes by crossing segment A is proportional to $P(A)$. These handoffs will be handled by the mobility agent (MA_{21}) in the common anchor cloud 2.1 (CAC_{21}).

As a second example, let segment B be $x_2 \leq x \leq x_4$ and $y = y_3$, where (x_2, y_3) and (x_4, y_3) lie on a circle. Let's assume that segment B is the only segment that belongs to the cloud 3.1. We formulate the probability of the mobile user crossing that segment as,

$$P(B) = \int_{x_2}^{x_4} \frac{1}{V_a} F_a(x, y_3) \mathrm{d}x \, \Delta y, \tag{6.13}$$

where Δy indicates a small positive thickness for the fine strip. The number of handoffs that the mobile user makes by crossing segment B is proportional to $P(B)$. These hanndoffs will be handled by the mobility agent in the cloud 3.1, which will thus become a common anchor cloud.

The thickness of segment A (Δx) and the thickness of segment B (Δy) are the same. The actual value for the thickness of Δx or Δy is not important in calculating the number of handoffs because the Δx and Δy cancels each other in Eqn. (6.14).

6.1.3 Formulation of Number of Handoffs

As stated in Section 5.1, $HO(MA)$ is the number of the handoffs handled by the mobility agent (MA) in the common anchor cloud. The calculation of $HO(MA)$ is as follows:

$$HO(MA) = \frac{P(S_{MA})}{\sum_{MA' \in (B_{in}(d_x) \cup B_{out}(d_x))} P(S_{MA'})} \cdot h \cdot t, \tag{6.14}$$

where the percentage of the handoffs handled by the mobility agent is multiplied with the mobile user's total number of the handoffs. The segment that belongs to the mobility agent is denoted by S_{MA}. $P(S_{MA})$ is proportional to the number of the handoffs that are

[8] Subscripts of x and y are used to differentiate the variables. They don't convey any special meaning.

handled by the mobility agent. The divisor is the probability of the mobile user crossing the strips that belong to all the mobility agents compared to anywhere else in the roaming area.[9] The set of the mobility agents that handle the mobile user's handoffs is presented with $B_{in}(d_x) \cup B_{out}(d_x)$.[10] We calculate the percentage of the handoffs handled by the mobility agent $\left(\frac{P(S_{MA})}{\sum_{MA' \in (B_{in}(d_x) \cup B_{out}(d_x))} P(S_{MA'})} \right)$. The total number of mobile user handoffs across the complete roaming area is calculated by $h \cdot t$ where h is the overall mobile user's handoff rate. A mobile user with higher h has higher velocity and makes more frequent handoffs in comparison to a mobile user with lower h.[11] The unit of h is number of handoffs per day. The variable t is the mobile user's total amount of activity time, which can be a week, month, or year, etc. We will assume t to be thirty days without loss of generality.

In Fig. 6.5, we consider only two segments, A and B, which belong to the common anchor clouds 2.1 and 3.1, respectively. All mobile user handoffs are due to crossing either of these segments. If the mobile user crosses segment A, then that handoff is handled by the mobility agent in the common anchor cloud 2.1. We calculate the number of handoffs handled by MA_{21} as:

$$HO(MA_{21}) = \frac{P(A)}{P(A) + P(B)} \cdot h \cdot t. \tag{6.15}$$

In the second example, when the mobile user crosses segment B, the handoff is handled by the mobility agent in the common anchor cloud 3.1 in Fig. 6.4. The mobility agent handling the handoff is denoted as MA_{31}. Then, the number of the handoffs is:

$$HO(MA_{31}) = \frac{P(B)}{P(A) + P(B)} \cdot h \cdot t. \tag{6.16}$$

Assuming that the mobile user's roaming area only intersects with segments A and B, the total number of the handoffs that the mobile user makes is the sum of $HO(MA_{21})$ and $HO(MA_{31})$.

In calculating the mobile user's handoff cost, the handoffs that the mobile user makes are categorized as in-domain and out-of-domain handoffs.[12] In-domain handoffs are

[9] As seen in Fig.6.4 and 6.5, there are empty (segment-free) areas in the roaming region of the user. The mobile user being on those areas do not result in any handoff.

[10] See Section 5.1 for set definition.

[11] We assume that a user's handoff rate (h) is available to the user. This is a realistic assumption, because a mobile device is capable of keeping track of handoffs. The experimental roaming data available in the Internet belong to a range of a conference venue which is limited [112]. Handoff data might also be obtained from simulation studies.

[12] Note that VMD can be deployed to any tier in the FCT internetworking model.

those that result when the mobile user crosses the strips within the initially registered VMD. Out-of-domain handoffs are those that result when the mobile user crosses the strips that are not within the initially registered VMD, but are within the mobile user's roaming region. The probability of mobile user crossing a strip can be found by using Eqn. (6.12) or (6.13) once the coordinates of the strips known.

6.1.4 Validation of Number of Handoffs

We obtain theoretical values for $HO(MA_{21})$ and $HO(MA_{31})$ from Eqns. (6.15) and (6.16), respectively. We aim to confirm the theorethical results with Monte Carlo simulations [145]. We run 100 Monte Carlo simulations and at each simulation we create 100,000 points that denote the locations a user visits.[13] [14] Fig. 6.6 shows the histogram of the randomly generated points that denote the locations a user visits. In Fig. 6.6, we observe the most points around $r = 4$. The reason for such a distribution is due to the 2D area affect. To illustrate, let's consider a small disc of radius 0.1 at the origin in the 2D plane. The small disc has an area of 0.01π. Now let's consider an annulus of inner radius 4 and outer radius 4.1. The area of the annulus is 0.81π that is much greater than the calculated area for the disc of radius 0.1 at the origin. Due to this 2D area effect, we see the numbers are more crowded around $r = 4$ in Fig. 6.6. Intuitively, there can be only a few points located at the origin due to the small size of the area, while more points can be located at greater distance from the origin. The density of the points will be decreasing towards the edge of the roaming area. From the mobile user's mobility perspective, it means that the mobile user has a few places he visits most frequently such as home and school, but there are more locations farther away from home that the mobile user visits less frequently.

In the roaming region, we also set up segments A and B as in Fig. 6.5 where $x_1 = 6$, $x_2 = -6$, $x_4 = 6$, $y_1 = -8$, $y_2 = 8$, and $y_3 = 8$. We take the handoffs per day, h, to be 10 and the number of days, t, to be 30. These numbers are picked for illustrative purposes only, as the same values will be used in both the analytical study and

[13] We compared the analytical and simulated mean values of the handoff amounts HO(MA21) and HO(MA31). We observe that the difference between the the simulation results and calculated analytical results was very close, to be precise within 0.01% error vicinity, so we found this configuration reasonable to run our simulations.

[14] We first divided the radius in 100,000 segments to decrease the discreteness (to make it closer to the continuous case). For example, for a 10 miles radius of a roaming region, Δx is 0.0001 miles. However, we create 100,000 points at each run due to computational limitations. In order to have a statistically averaged number of points fall into each segment, number of points should be much greater than the number of segments. On the other hand, reducing the number of segments decreases the resolution of the experiment as a trade off. Therefore, we select a new Δx which corresponds to the thickness of 3,000 segments to downconvert the number of segments to approximately 33, such that we acquire enough statistical average and a decent resolution at the same time.

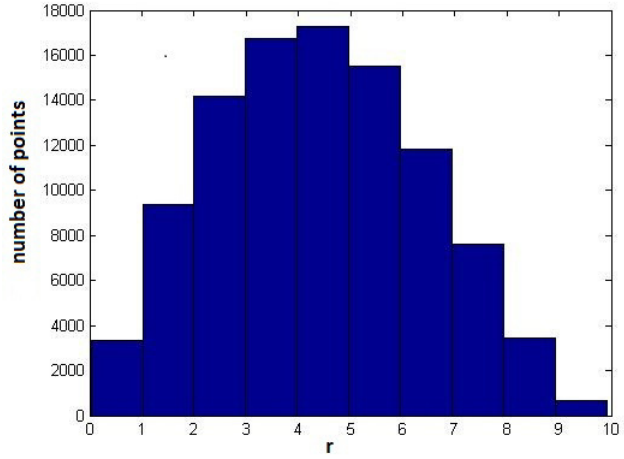


Figure 6.6: Histogram of the randomly generated 100,000 points as function of r.

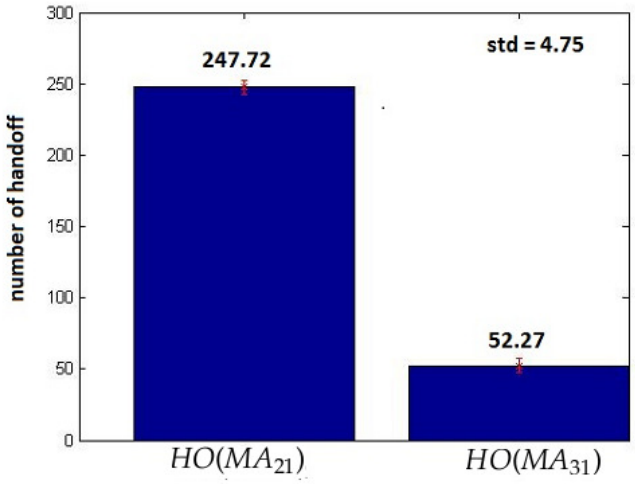


Figure 6.7: The number of handoffs handled by MA_{21} and MA_{31} in the Monte Carlo simulations.

the Monte Carlo simulations; the choice of values is not important to achieve our goal. We separately count the number of the random points that fall onto segments A and B. Then, we calculate $HO(MA_{21})$, the number of handoffs handled by MA_{21}, by multiplying the total number of handoffs, which is given by $h \cdot t$, by the ratio between the number of points falling on segment A and the total number of points falling on either segments A

or B. We calculate $HO(MA_{31})$, the number of handoffs handled by MA_{31}, similarly. We run this Monte Carlo simulations hundreds of times and present the results in Fig. 6.7. The mean of $HO(MA_{21})$ is 247.72, while the mean of $HO(MA_{31})$ is 52.27 with a standard deviation 4.75 - for both of them.

We also calculate $HO(MA_{21})$ and $HO(MA_{31})$ by using Eqns. (6.15) and (6.16). In our analytical calculations, we use the same coordinates for segment A and B as in Fig. 6.5. Edge coordinates of segment A are $x_1 = 6$, $y_1 = -8$, and $y_2 = 8$, while the end points of segment B are $y_3 = 8$, $x_2 = -6$, and $x_4 = 6$. The radius of the roaming area is 10 miles. We take the handoffs per day, h, to be 10 and the number of days, t, to be 30. We obtain 247.83 and 52.16 for $HO(MA_{21})$ and $HO(MA_{31})$, respectively. Considering the standard deviation in the Monte Carlo study, the simulated results agree well with the theoretical results. In conclusion, we confirm the theoretical results with the results obtained from the Monte Carlo simulations.

In the previous sections, we provided a guideline for obtaining the number of handoffs of a mobile user whose movements follows the proposed mobility model. In general, mobile users have various mobility profiles, QoS preferences, handoff cost sensitivities. We model a handoff cost optimization problem in the next section. Our goal is to find an optimum domain that aligns with mobile users' preferences and also incurs minimum handoff cost mobility preference.

6.2 Handoff Cost in VMD

As explained in the Introduction section, mobile users have various mobility preferences and cost concerns. Mobile users would like to connect to a VMD that fits their mobility characteristics, handoff cost concerns, and delay requirements. Depending on the network parameters and a mobile user's preferences, the VMD that a mobile user is willing to register with may change. Therefore, in this section, we will model the handoff cost function as an optimization problem to find out the optimum VMD that a mobile user should register with.

We want to identify the type of the handoff cost optimization problem in order to decide which method to apply for a solution. For this purpose, we will analyze the decision variables, parameters, and the constraints in the handoff cost optimization problem. The findings will help us decide which numerical method to apply in finding the optimum VMD for mobile users in the next section.

6.2.1 VMD Handoff Cost Optimization Problem

The following equation formulates the handoff cost of a mobile user for a given VMD d_{vmd} that a mobile user is willing to connect to as

$$H(d_{vmd}) = w_p \cdot (\mu \cdot Sto(d_{vmd}) + \theta \cdot Sig(d_{vmd})) + w_d \cdot Data\_Loss(d_{vmd}), \tag{6.17}$$

which comprises (i) the storage cost at proxy AAA servers and the forwarding bases in a VMD ($Sto(d_{vmd})$, formulated in Eqn. (5.4)); (ii) the signaling cost that is incurred due to handoff support ($Sig(d_{vmd})$, calculated in Eqn. (5.5)); and (iii) the cost of data loss due to handoff latency ($Data\_Loss(d_{vmd})$, derived in Eqn. (5.9)). These costs are in bytes. Details of these costs components are in Section 5.1 and 5.2.

Using the handoff cost function, we aim to find the optimum VMD that is denoted with d^*_{vmd} that a mobile user should register with. The optimum VMD ensures the minimum handoff cost defined by Eqn. (6.17) for the given mobile user preferences. Therefore, in the optimization study, $H(d_{vmd})$ is our decision function, and d_{vmd} is our decision variable. While finding the optimum VMD, we are actually interested in finding the tier of VMD in the topology. As explained in Section 6.1.1, our topology is a balanced tree and hence the VMDs have the same number of child clouds and coverage area size as long as the tier of the cloud that they are rooted to are the same. For the given mobility characteristics of a mobile user and network conditions, the tier of d^*_{vmd}, $T(d^*_{vmd})$, will be an integer such that:

$$1 \leq T(d^*_{vmd}) \leq K, \quad \forall T(d^*_{vmd}) \in \mathbb{Z}^+ \tag{6.18}$$

where $K = 6$ [2]. There is only one variable in our handoff cost function and that is an integer; hence, our optimization problem is an integer problem.

The parameters in our handoff cost function are w_d, w_p, μ, τ, θ, λ, and R_{data}. See Section 5.1 for details. These parameters are retrieved from the mobile user and the service provider. In the handoff cost function, w_d denotes a mobile user's relative sensitivity to the cost of data loss due to the handoff and w_p denotes the mobile user's relative sensitivity to the storage and the signaling costs. The sum of these parameters has to be as follows:

$$w_p + w_d = 1. \tag{6.19}$$

Users may have various sensitivity to data loss costs or service-provider costs depending on their preferences and needs. Therefore, the ranges of the values that w_d and w_p can get are the following:

$$0 \leq w_d \leq 1, \tag{6.20}$$

and

$$0 \leq w_p \leq 1. \tag{6.21}$$

Further, the relative impact of storage and signaling costs is provided with weights μ and θ parameters, respectively. The values of these weights can be in the following ranges - similar to w_p and w_d:

$$0 \leq \mu \leq 1, \tag{6.22}$$

$$0 \leq \theta \leq 1, \tag{6.23}$$

and

$$\mu + \theta = 1. \tag{6.24}$$

The cost multiplier τ denotes an extra cost of getting services from a service provider that serves a user temporarily as the user roams into its service area.[15] It gets its values as follows:

$$\tau_{min} \leq \tau \leq \tau_{max}. \tag{6.25}$$

The minimum and maximum values of τ, τ_{min}, and τ_{max}, respectively, depending on the agreement between the initially registered service provider and the temporarily registered service provider.

Data usage of a mobile user is denoted as $\lambda_s \cdot R_{data}$ in unit of bytes per second, where λ_s is the number of sessions that the mobile user has per unit of time and R_{data} denotes the average number of data bytes per session. The range of data usage is as follows:

$$data\_usage_{min} \leq \lambda_s \cdot R_{data} \leq data\_usage_{max} \tag{6.26}$$

where the minimum and maximum data usage values ($data\_usage_{min}$ and $data\_usage_{max}$ respectively) depend on the mobile user's characteristics and the underlying network limits such as bandwidth.[16]

Table 6.1 shows the system constants that depend on the FCT internetworking model. We pick the following values to be consistent with our previous simulation and analytical studies [2,140–142] in which we aim to represent a real network condition [See Tables 4.1 and 5.1 for details].

[15] τ was first introduced in Eqn. (5.7).

[16] One might think that minimum data usage is naturally zero. So, there is no need to define a lower limit on data usage. However, there can be scenarios where the mobile device applications running at the background also consuming Internet data. The maximum data usage is necessary because of the limitations of the underlying technologies. For example, a mobile user can not have 1 terabytes per second data usage with his mobile phone using the current network technologies.

Table 6.1: The VMD system parameters.

Parameter	Value	Unit
γ, the number of children of a parent in the FCT internetworking model	4	unitless
η, the cost of storing mobile user profile	250	bytes
δ, the cost of storing forwarding information	10	bytes
B_{wl}, the bandwidth of a wireless link	100	megabit per second
B_w, the bandwidth of a wired link	1	gigabit
L_{wl}, the propagation delay of a wireless link	0.002	seconds
L_w, the propagation delay of a wired link	0.01	seconds
p_q, the average processing and queuing time of a packet at a router	0.1	seconds
m, the size of a mobility control message	$76 \leq m \leq 326$	bytes

The handoff cost function is nonlinear because the data loss cost function Eqn. (5.9) and the signaling cost function Eqn. (5.5) are nonlinear. The data loss cost function has the multiplication of two components both of which depend on the decision variable d_{vmd}. These two components are the number of in-domain/out-of-domain handoffs ($HO(MA)$) and delays ($D_{in}(MA)$ / $D_{out}(MA)$). The number of in-domain and out-of-domain handoffs depends on where MA is. MA can be a member of either $B_{in}(d_{vmd})$ or $B_{out}(d_{vmd})$ depending on d_{vmd} [See Section 5.1 for details.]. The decision variable d_{vmd} also decides if the handoff is in-domain or out-of-domain. The in-domain and out-of-domain handoff delays also depend on d_{vmd} as expressed in Eqns. (5.31) and (5.33). Further, the number of items to be summed in the data loss cost function depends on the size of $B_{in}(d_{vmd})$ and $B_{out}(d_{vmd})$, both of which depend on d_{vmd} as stated before. Due to all these dependencies of data loss cost components on d_{vmd} the data loss cost function is nonlinear.

The signaling cost function Eqn. (5.5) has the multiplication of a number of in-domain/out-of-domain handoffs ($HO(MA)$) and in-domain/out-of-domain signaling costs ($C_{in}(MA)$ / $C_{out}(MA)$) - similar to the data loss cost function. The multiplier and the multiplicand values depend on d_{vmd}. The MA can be in d_{vmd} or out-of d_{vmd}. So the handoffs that are handled by the MA can be either in-domain or out-of-domain depending on d_{vmd}. The in-domain and out-of-domain handoffs incur different in-domain/out-of-domain signaling costs, as expressed in Eqns. (5.29) and (5.30), respectively. Further, the number of items summed in the signaling cost function is decided by the number of MAs in the sets of $B_{in}(d_{vmd})$ and $B_{out}(d_{vmd})$, which depend on d_{vmd}. Due to all these dependencies on d_{vmd} the signaling cost function is nonlinear. The nonlinearity of the handoff cost function can be observed in Fig. 5.3 and 5.8 among others. As illustrated in the figures, the handoff cost value does not change linearly with the tier of the initially registered VMD.

In summary, our handoff cost optimization problem has an integer variable and a nonlinear objective function (i.e., handoff cost function). We used a mixed-integer, non-

linear programming (MINLP) solver [146] because the characteristics of our handoff cost optimization function is within the set of the problem types that MINLP solver can solve. MINLP solves nonlinear problems that has either integer or continues decision variables.

6.3 Numerical Study

We create mobile user profiles who follow the proposed mobility model. The mobile user profiles differ in terms of roaming area range and origin, and handoff frequencies. We would like to demonstrate how these differences affect the number of handoffs and the distribution of the handoffs to the mobility agents. As part of the numerical optimization study, we create mobile user profiles with different cost sensitivities (w_p and w_d), and data usage ($\lambda \cdot R_{data}$). We consider service providers that enforce different external service cost multipliers (τ). We would like to demonstrate that optimization results are in harmony with the proposed optimization model. We assign empirically realistic values to each parameter (see Tables 4.1 and 5.1) to be consistent with the numerical studies in the previous chapters, [2, 140–142], and to reflect the current network technologies.

6.3.1 Numerical Study of the Mobility Model

We discuss the proposed mobility model according to the numerical results. We have created several mobile user profiles with the same mobility model but with different roaming area ranges and origin, and handoff frequencies. We set the number of children of a cloud, γ, to 4 to have symmetric distribution of the clouds on the deployment area and to be consistent with our previous studies in Section 4. The FCT internetworking model is mapped on a square area, which has a side length of 256 miles, as illustrated in Fig. 6.4. The side length is set to a value that is a power of γ to have integer values for the strip coordinates and to create an easy illustration.

Roaming Range of Mobile Users

In our mobility function, $m_a(r)$, we represent the roaming range of a mobile user with a where it decides on the radius of the roaming region of a mobile user such that $0 \leq r \leq a$ in Section 6.1. To illustrate, we have created three mobile user profiles:

- *User 1* has a short roaming range with a radius of 16 miles, which means $a = 16$. User 1 is regarded as a local mobile user.
- *User 2* has a medium roaming range with a radius of 32 miles, which means $a = 32$.

- *User 3* has a wide roaming range with a radius of 64 miles, which means $a = 64$.

We set the origin of the mobile users (x_0, y_0) to $(64, 64)$, which is the center of the first quadrant square on the area where the FCT internetworking model is mapped in Fig. 6.4. Locating the mobile users on other quadrants will not change our results because those squares are symmetrical to each other and have the same number of strip segments. Note that we chose these values for illustrative purposes, and the program allows mobile users to be located anywhere in the area. We set the values for h to 10 handoffs per day and t to 30 days - again for illustrative purposes. Our aim is to find the distribution of handoffs to the VMD tiers.

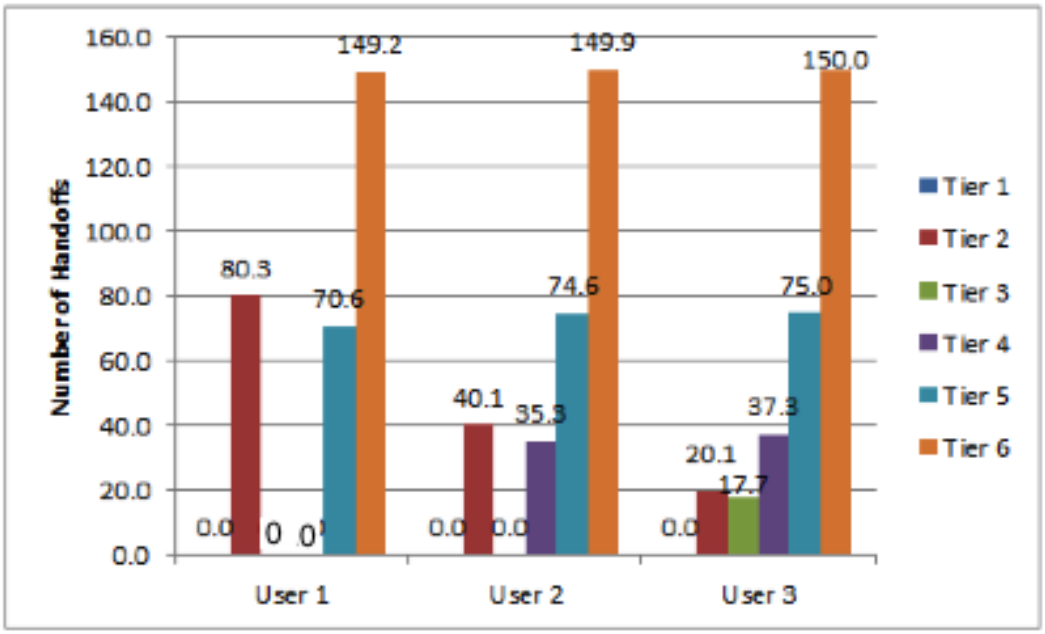


Figure 6.8: The number of the handoffs of User 1, User 2, and User 3.

Fig. 6.8 presents the total number of the handoffs by the aforementioned mobile users along with the tiers of the common anchor cloud. We retrieved these handoff values using Eqn. (6.14). For instance, User 1 does 80.3 handoffs, handled by mobility agents in tier-2 common anchor clouds; 70.6 handoffs, handled by mobility agents in tier-5 common anchor clouds; and 149.2 handoffs, handled by the mobility agent in tier-6 common anchor clouds. User 1 does not make any handoff handled by mobility agents in tier-1, tier-3, or tier-4 common anchor clouds because the mobile user's roaming region does not intersect with any of these common anchor clouds' boundary strips. User 2 does not cross any

boundary strip belonging to tier-1 or tier-3 common anchor clouds, while User 3 does not cross any boundary strip belonging to the tier-1 common anchor cloud. The expansion of the roaming region results in more handoffs on the different tiers, which affects the distribution of the handoffs to the tiers. In the case of User 3, we observe that the number of handoffs handled by the tier-2 common anchor clouds is more than the number of the handoffs handled by the tier-3 common anchor clouds, because the roaming region of User 3 involves more tier-2 strips than tier-3 strips, which is due to the roaming region of the mobile user and the FCT internetworking model deployment. User 1, User 2, and User 3 do the same number of the handoffs in total, because they do the same number of handoffs per day (h), and their activity time (t) is the same. The range of the roaming area of the mobile users does not affect the number of the handoffs, because the number of handoffs is calculated by $h \cdot t$.

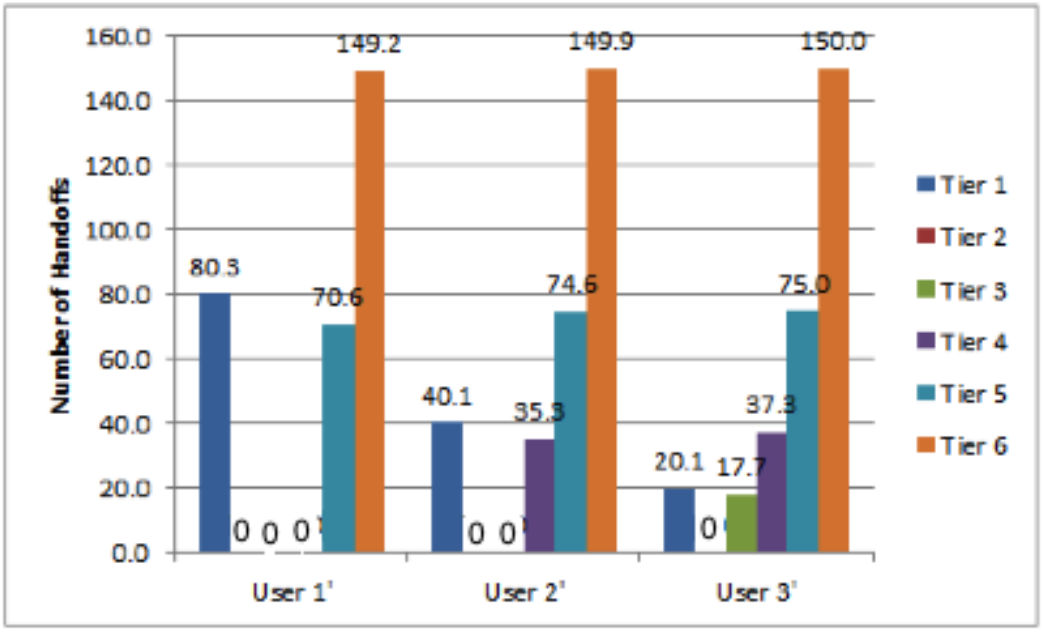


Figure 6.9: The number of the handoffs of User 1′, User 2′, and User 3′.

We created three more mobile user profiles: User 1′, User 2′, and User 3′, which differ from the previous mobile user profiles only in the coordinates of their origins, which are now (0,0). We aim to analyze the effect of the origin on the distribution of the number of the handoffs to the tiers. The results are presented in Fig. 6.9. The User 1′, User 2′, and User 3′ now have tier-1 handoffs instead of tier-2 handoffs, because their roaming region intersects with the tier-1 common anchor cloud's boundary strips instead of the

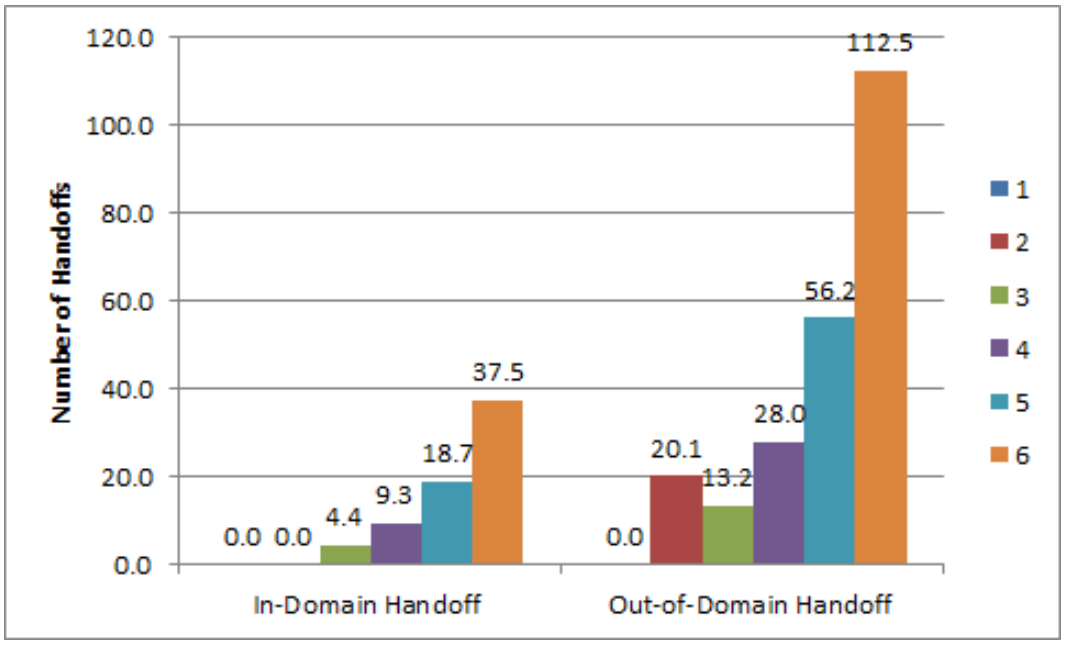


Figure 6.10: The distribution of the in-domain and out-of-domain handoffs of User 3.

tier-2 common anchor cloud's boundary strips. The number of the tier-1 handoff of User 1', User 2', and User 3' are the same with the number of the tier-2 handoff of User 1, User 2, and User 3, respectively, due to having the same ranges of roaming regions and covering the same number of strip segments. We observe that the number of handoffs in the other tiers does not change because the location of the segments relative to the center of the roaming region and the length of the segments are the same for all the users. The reason of this distribution of the segments is due to the mapping algorithm that is presented in Section 6.1.1. In addition, all the users move according to the same proposed mobility model, hence, the probability of mobile users' crossing those segments are the same. Therefore, we have the same distribution of he handoffs to the tiers.

The mobile user's handoffs will be either in-domain or out-of-domain depending on the VMD tier to which the mobile user is initially registered. To illustrate, in Fig. 6.10, we present the handoffs done by User 3, in the case that the mobile user is initially registered to the VMD at tier 3, in which case the mobile user will make both in-domain and out-of-domain handoffs. As seen in the Fig. 6.10, User 3 does out-of-domain handoffs handled by mobility agents in common clouds at various tiers because the mobile user's roaming region exceeds the coverage area of the initially registered VMD. We observe that a number of in-domain handoffs handled at each tier is less than the number of the

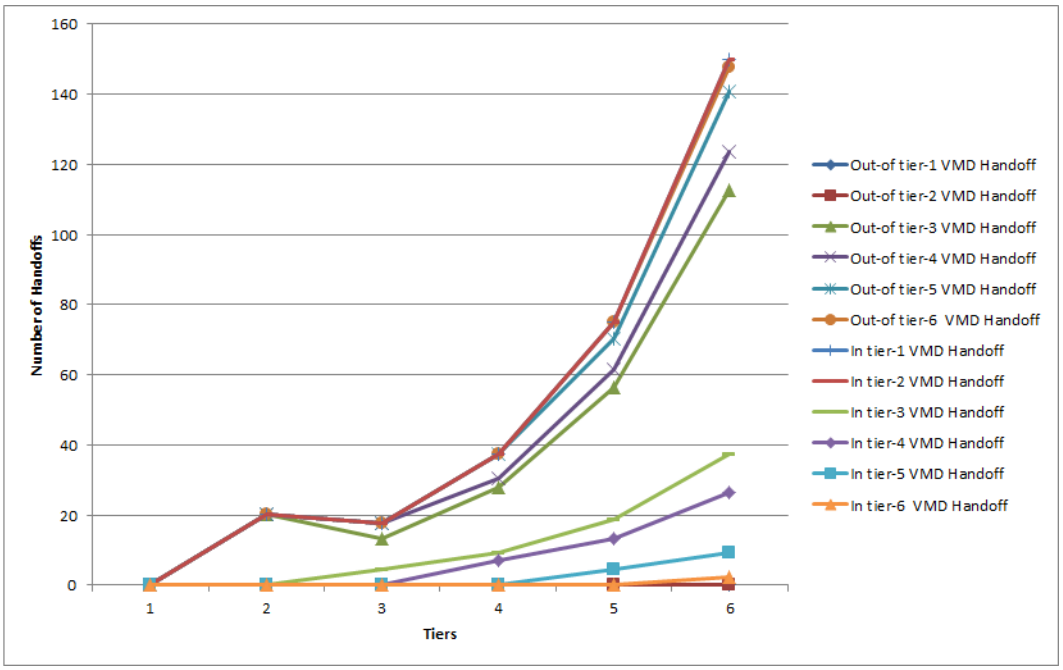


Figure 6.11: The number of handoffs as a function of the tiers of the common anchor clouds for the cases when User 3 is registered to the VMDs at tier 1 - 6.

out-of-domain handoffs handled at each tier, because the mobile user is registered to the VMD at tier 3, which covers only a small portion of the area in which the mobile user roams. The number of the out-of-domain handoffs increases as the tier value increases because of the increasing number of common anchor clouds that the VMD cannot cover.

In Fig. 6.11, we present User 3's in-domain and out-of-domain handoffs when the mobile user is registered to the VMD at tiers 1 to 6. The mobile user can do either in-domain or out-of-domain handoffs, while he is registered to a VMD at tiers 1 to 6. To represent different combinations of handoffs with initially registered VMDs, we name each combination as follows. For example, we used "out-of tier-1 VMD handoff " to denote the out-of domain handoffs for the case when the mobile user is initially registered to the VMD at tier 1. The x axis has the tier of the mobility agents that handle the handoffs. When the mobile user is registered to the VMD at tier 1 or 2, all of the mobile user's handoffs are in-domain handoffs, because these VMDs cover the whole roaming region of the user. For the other registered VMDs, as the tier of the VMD increases, the number of in-domain handoffs decreases, and the number of out-of-domain handoffs increases, because the VMDs' coverage areas are narrowed. As seen in Fig. 6.11, the number of handoffs decrease at tier 3, because the mobile user's roaming region intersects with a

smaller area of the strips that belong to tier 3 common anchor clouds, compared to the area of the strips that belong to tier 2 and tier 4 common anchor clouds. As expressed in Eqn. (6.12), the probability of a mobile user crossing the segment is proportional to the area of that segment.

6.3.2 Numerical Study of the Handoff Cost Optimization

In this section, we aim to solve the handoff cost problem for a given mobile user 's mobility preferences using DICOPT [147], which is an MINLP solver in GAMS [144]. Then, we illustrate that the optimization results are in harmony with the proposed optimization model in Section 6.2. We consider mobile users with different cost sensitivities (w_p and w_d), data usage ($\lambda \cdot R_{data}$), and service providers that enforce different external service cost multipliers (τ). We used mobility profile of User 3 to calculate the number of handoffs. See Section 6.3.1 for details. The reason for using User 3 is that this mobility profile has handoffs that are handled by mobility agents at a variety of tiers compared to other profiles. Therefore, the optimum domain could be any domain depending on the values of w_p, w_d, $\lambda \cdot R_{data}$ and τ.

Effect of User's Cost Sensitivity Parameters

In this section, we aim to observe the effect of sensitivity to costs related to handoff latency (w_p) and the ones imposed by the service provider (w_d). We have studied mobile users who have different w_p and w_d values and observe their handoff costs and optimum VMD tier. All the other parameters are the same: $\mu = 0.5$, $\theta = 0.5$, $\tau = 1.1$, and $\lambda_s \cdot R_{data} = 10240$. Note that the parameter values do not change our handoff cost framework or optimization mechanism. We give the same values to μ and θ, because we do not want to differentiate the storage and signaling costs. The τ value is assigned, considering that the cost receiving service from a temporarily registered service provider is higher than the permanent service provider. We assume that the mobile user has 10240 KBps data usage during handoffs considering the wireless access technologies [148] [149] to be empirically realistic. These values are used in our previous studies (see Section 5 for details) to represent a typical mobile user and his/her preferences. We would like to be consistent with our choices throughout this work.

In Fig. 6.12, we draw how storage, signaling, data loss and total costs are impacted from varying cost sensitivity pairs (w_p and w_d) on the x axis. Each cost sensitivity pair belongs to a mobile user. We do not consider each possible value of cost sensitivity pairs. For example, we consider the mobile users with cost sensitivities (w_p=0.1 & w_d=0.9), (w_p=0.2 & w_d=0.8) but not w_p=0.15 & w_d=0.85, because Fig. 6.12 clearly show that the cost sensitivities linearly impact the handoff cost components. The linearity between cost sensitivities and costs are also clear in Eqn. (6.17). A mobile user with higher w_p

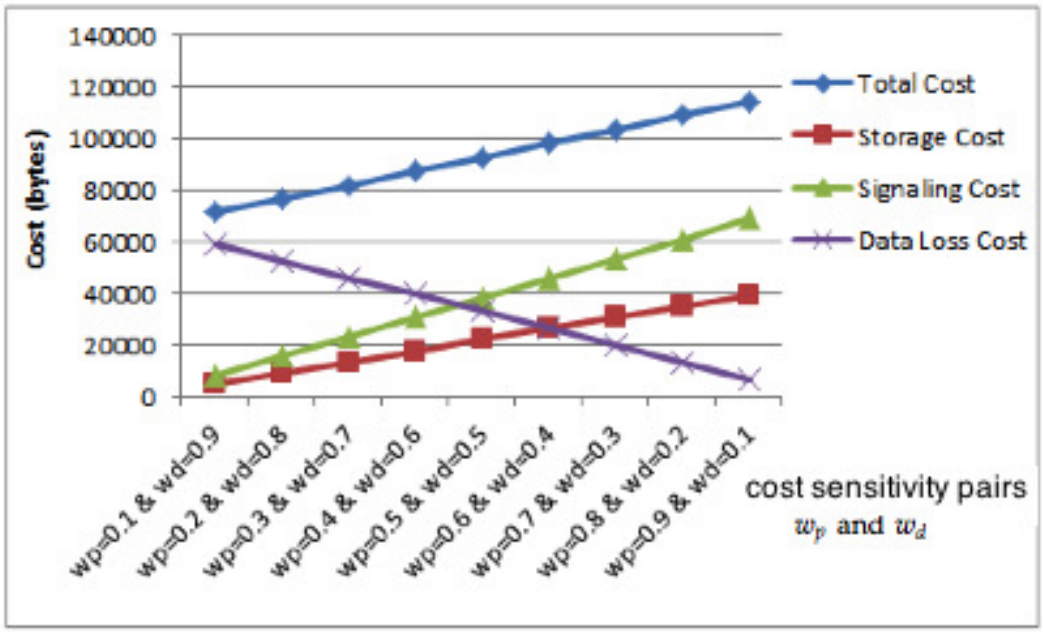


Figure 6.12: The effect of w_p and w_d on storage, signaling, data loss, and total handoff costs.

has a higher signaling and storage costs and vice versa. The mobile user with lower w_d observed lower data loss cost and vice versa. Further, we observe that the optimum VMD tier is tier 3 for all mobile users, because with a given mobility profile and preferences, VMDs at other tiers are still costlier than VMDs at tier 3.

Effect of External Service Cost Parameter

In this section, we discuss how external service cost multiplier (τ) affects the handoff cost. In Fig. 6.13, we illustrate the storage, signaling, data loss, and total handoff costs for the cases that the mobile user temporarily receives services from different service providers that impose different external service cost multipliers. We present the tier of the optimum domain in Fig. 6.14 for given external service cost multipliers.[17] The values of the external service cost multipliers range from 0.2 to 2. We did not use lower values of τ because the optimum domain for ($\tau = 0.2$) was already the domain at the lowest tier, that is tier 6. Note that the optimum VMD tier for other τ values (with 0.1 step increase in the range between 0.2 and 2) is the same as the VMD tier of the preceding τ. We omit these τ values

[17]User 3's preferences stay the same: $w_p = 0.3$, $w_d = 0.7$, $\mu = 0.5$, $\theta = 0.5$, and $\lambda_s \cdot R_{data} = 10240$.

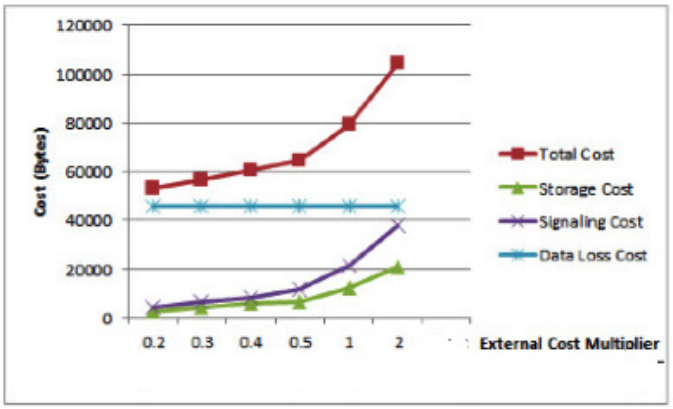


Figure 6.13: The effect of external cost multiplier, τ, on storage, signaling, data loss, and total handoff costs.

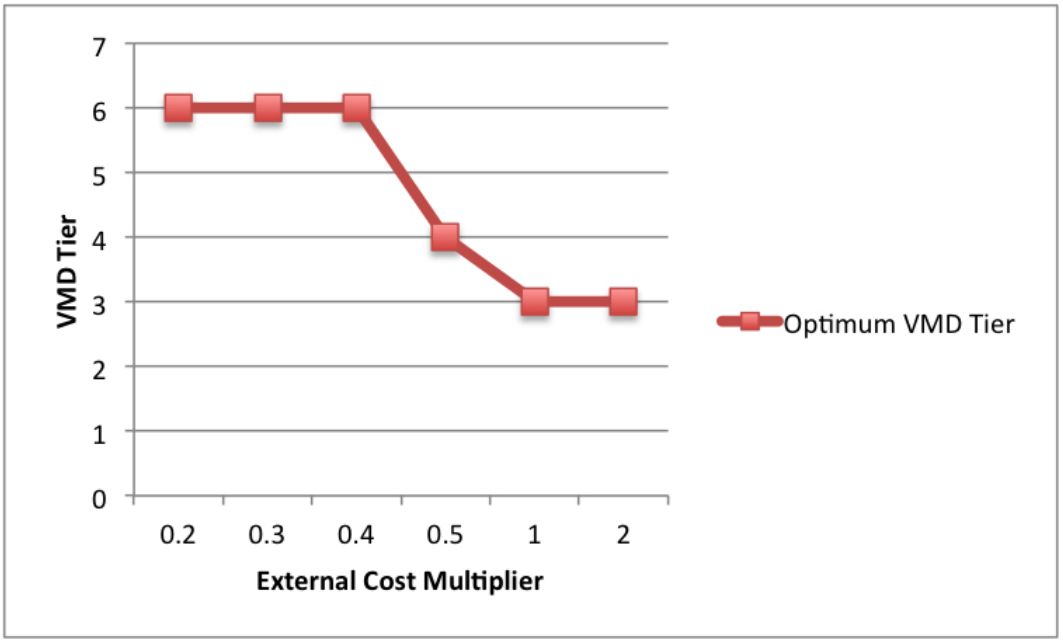


Figure 6.14: The effect of external cost multiplier, τ, on the optimum VMD tier.

to provide clear presentation.

In the GAMS output, we observe that the optimum VMD tier goes up towards tier 3 as the τ values increase because User 3 needs to limit the service consumed from tem-

porarily registered VMD as the service becomes costlier. User 3 limits the consumption of the service by registering to a wider VMD, hence, decreasing the out-of-domain handoffs. Signaling cost, data loss cost, and eventually total handoff cost increases with the increasing value of τ, because User 3 performs out-of-domain handoffs, which are served by the temporarily registered VMD. Data loss cost does not change as it is purely related to the handoff latency and does not depend on τ as expressed in Section 5.2.3.

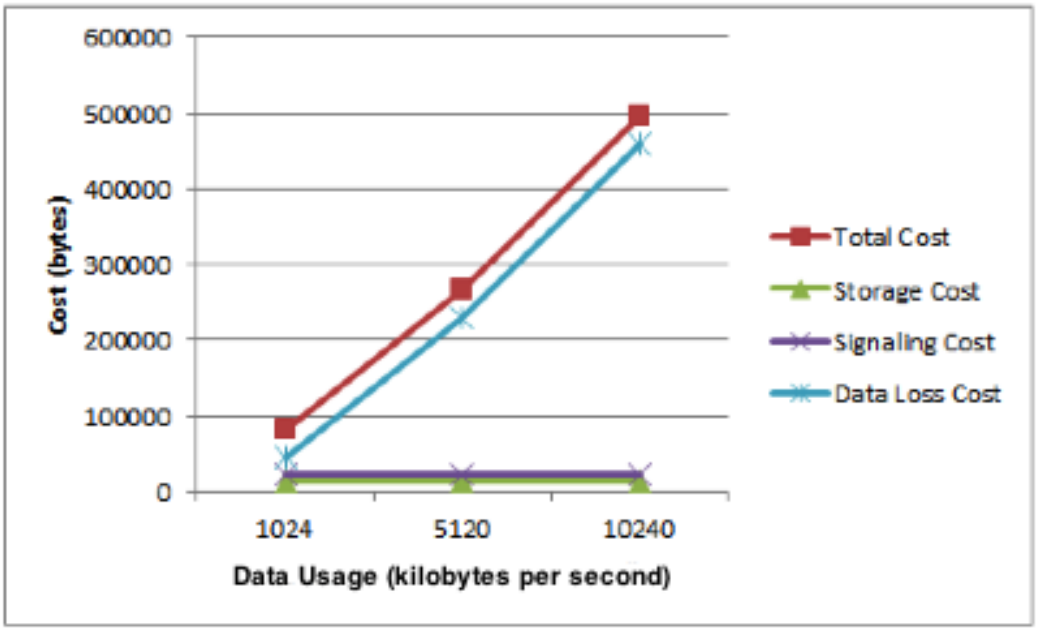


Figure 6.15: The effect of $\lambda_s \cdot R_{data}$ on storage, signaling, data loss, and total handoff costs.

Effect of User Data Usage Parameter

We discuss the impact of a mobile user's data usage character on the handoff cost in this section. We created three mobile user profiles who follow User 3 mobility preferences. These three mobile user profiles differ in terms of the data usage ($\lambda_s \cdot R_{data}$) characteristics.[18] In Fig. 6.15, we present the storage, signaling, data loss, and handoff cost values for each data usage. We observe that the data loss cost is higher for the mobile users with higher data usage. The data loss cost is higher because there will be more data loss per unit time during the handoff. We calculate the data loss cost by multiplying the handoff

[18]We used only three different values of data usage because the linearity between the data usage and the data loss cost was clear in the results. Having more values would not change the observation of the linearity.

latency with the data usage as expressed in Eqn. (5.9). On the other hand, the signaling and storage costs stay the same as they are not dependent on the mobile user's data usage.

6.4 Summary

In this chapter, we have studied the handoff cost optimization problem. We identified one decision variable that was an integer and the objective function was nonlinear. Consequently, the optimization problem can be solved with integer nonlinear programming tools. Solving the optimization problem allowed us to find out the optimum VMD tier, which resulted in the minimum handoff cost for a given mobile user's mobility preferences.

To create an optimization model, we provided a mechanism to find the number of handoffs from any given mobility model of a mobile user. We introduced a mobility function to represent the movement of a mobile user who roams mostly in the center of the roaming area and less toward the edges of the roaming area. The mobility function enabled us to ascertain a mobile user's frequency of presence in any part of a roaming area that has a radius of a. We presented the results from the same mobility function with different inputs such as varying roaming range, roaming area origin, and handoff frequency. We examined the results of the numerical mobility study and concluded that they are in accordance with the theory and our intuition.

We study mobile users with different cost sensitivity, data usage and with the service providers that impose different external service cost multipliers to find out the change on the handoff cost and the optimum VMD with which the mobile user should register. The change in the parameter values affected the related cost components linearly. The results are in accordance with the analytical models.

Chapter 7

Conclusions

This chapter provides a summary of this dissertation work and the direction of future work. The dissertation began in Chapter 1 by introducing the motivation of this Ph.D. study for the need of a seamless mobility service that can be provided by the clean-slate future Internet architectures. We present the research challenges in mobility management and explain our proposed novel solution that is called the Virtual Mobility Domain Architecture.

Chapter 2 surveys the current and future Internet protocols over the period of 2002-2012. The identity and handoff-management design fundamentals of MIPv4, MIPv6, Fast MIPv6, HMIPv6, and PMIPv6 protocols are presented to better present the current challenges of the Internet. We then analyzed the identity and handoff-management methodologies of the next-generation mobility solutions supported by the future Internet initiatives around the world: MobilityFirst, XIA, Ambient Networks, DAIDALOS, AKARI, HIP, i3, Hi3, LISP, MILSA, CARMEN, HURRICANE, and MobileNAT. A qualitative comparison of the aforementioned protocols and schemes has been made by using a unified platform. We further present the most commonly used mobility models in the literature. Then, the previously conducted handoff-cost optimization studies are explained. The aim of this chapter is to provide related works for the research explained in the following chapters.

The design and implementation of a novel future Internet mobility architecture called Virtual Mobility Domain is presented in Chapter 3. The VMD proposes a novel addressing scheme with a unique address-acquisition technique, a network-based collaborative handoff-management scheme for intra-AS and inter-AS roaming needs of mobile users. The VMD is built to work on the Floating Cloud Tiered (FCT) internetworking model, which is derived from the tiered structure existing among ISP networks, to define their business and peering relationships. Leveraging the tiered structure and its hierarchical

properties, the VMD provides user-centric overlapping mobility domains, as it allows mobile users to register to VMDs of varying scopes that are suited to the users roaming needs.

In Chapter 4, the handoff performance of the VMD architecture is analytically modeled for signaling overhead and handoff latency metrics, which are main indicators of a seamless handoff service. The analytical and OPNET simulation-based performance analysis of the VMD for varying intra-AS deployment and across multiple-AS deployment has been conducted in comparison with MIPv6, HMIPv6, and PMIPv6. The performance results reveal that the VMD outperforms MIPv6, HMIPv6, and PMIPv6significantlyin terms of handoff latency, packet loss, and signaling overhead, which is mainly due to the network-based unique collaborative handoff-management scheme. The relative performance improvements achieved with the VMD will serve as a benchmark for future mobility-related studies.

The novel, handoff-cost framework that accounts for all metrics of interest of a mobile user, such as registration costs, latency in handoff, data loss during handoff, and signaling overhead, is proposed in Chapter 5. The unified platform helps the Internet-connected mobile users examine the registration costs, signaling overhead, and data-loss performance of any mobility protocol. To illustrate, we applied the framework to assess the performance of IPv6-based mobility protocols and the VMD architecture. We then employed the framework for a user to find the optimum VMD with which to register, leveraging that the VMD architecture is user-centric and allows a user to register to a VMD at any tier in the Internet architecture. We analyzed the effect of each mobility preference and cost sensitivities of a user, including service provision parameters of a system on the handoff-cost components and identified the optimum VMD with which the user should register.

In Chapter 6, we present the optimization of the handoff cost considering a mobile user as the primary focus. We first conduct a novel mobility study on providing a guide for finding the number of handoffs in a typical VMD. This study can be applied to any mobility model due to its generic framework. We leverage the output of this study during the handoff-cost-optimization study. The aim of a handoff-cost-optimization study is to find out the optimum VMD tier that results in the minimum handoff cost for the given user mobility preferences. We first identify that the decision variable is an integer and the nature of the objective function is non-linear. In that way, we show that the handoff-cost-optimization problem can be solved by integer, non-linear programming tools. We have modeled our handoff-cost-optimization problem on a general algebraic modeling system (GAMS), and then we leverage the DICOPT solver to solve our handoff-cost-optimization problem. We conduct a numerical study by varying the optimization parameters in a GAMS. The numerical results are in accordance with the intuition gained

from the analytical models.

The presented thesis work can be expanded to new future directions. We would like to provide a general discussion on future research directions that can be based on our work. The first future direction of the work is to deploy the VMD architecture on a general global environment for network innovations (GENI) test bed [9] where the FCT internetworking model is currently operating. It would be interesting to investigate the performance of the VMD in a real network environment and compare the performance with the analytical and simulation-based results. The second interesting direction of future work would be to investigate the integration of the VMD architecture built on the FCT internetworking model and cellular architectures, such as long-term evolution (LTE). The third direction of the future work would be to study the network economics based on the proposed framework and adding the service providers' preferences to create a game theoretic model where the players are the mobile users and the service providers. As the last direction of the future work, it would be interesting to study the secure registration and communication. The tiered architecture may enable collaborative security management by leveraging distributed hash tables.

Appendix A

The Category Definitions

The followings are the definitions of the categories at Tables 2.2, 2.3, 2.4, 2.5, and 4.2.
Mobility Scope: The type of a mobile node mobility that the protocol deals with. i.e. macro or micro mobility.
Mobility Management: Shows whether a mobile node is involved in mobility management related signaling or not. If the mobile node contributes to the part of the mobility management related activities, and then it is called host-based mobility management. Otherwise, network-based mobility management.
Network Architecture: If a new architecture or protocol provides hierarchy between the new network nodes and there is a collaboration on mobility management, then this architecture is called hierarchical. Otherwise, the architecture is called as flat.
Target Network: The type of network that the protocol aims to provide mobility management.
Operating Layer: The OSI layer(s) that the specified protocol operates on.
Required Infrastructure: New network nodes are required by the specified protocol or architecture to be able to operate properly.
Mobility Protocol: The name of the mobility protocol used in the specified scheme.
MN Modification: Whether the specified architecture or protocol requires modification on a mobile node or not to be able to operate properly.
MN Address: Presents the general name given to a mobile node's address.
Address Type: Presents the type of the mobile node's address.
Address Length: The length of the mobile node's address in bits.
Address Change: Whether the address received by a mobile node in the protocol domain changes or not when the mobile node moves to new network.
Address Assigned by: The name of the node or agent in the network that assigns an address to a mobile node.
Tunneling: Shows if tunneling used for the data communication between a mobile node and a correspondent node. If yes, then where the tunneled communication is applied,

inter-AS or intra-AS.

Appendix B

Validation of Probability Density Functions

We will validate the probability density functions $f_R(r)$ and $h(\theta)$ in Eqn. (6.10). As a validation of $f_R(r)$, we look to see $f_R(r)$ integrate to one:

$$\int_0^a \frac{1}{A} m_a(r) r \, \mathrm{d}r = \frac{1}{A} \cdot \int_0^a m_a(r) r \, \mathrm{d}r = \frac{1}{A} \cdot A = 1. \tag{B.1}$$

As a validation of $h(\theta)$, we look to see $h(\theta)$ integrate to one:

$$\int_0^{2\pi} h(\theta) d\theta = \int_0^{2\pi} \frac{1}{2\pi}(1) d\theta = \frac{1}{2\pi} \cdot \int_0^{2\pi} d\theta = \frac{1}{2\pi} \cdot 2\pi = 1. \tag{B.2}$$

Bibliography

[1] D. Johnson, C. Perkins, and J. Arkko, "Mobility Support in IPv6," RFC 3775 (Proposed Standard), Internet Engineering Task Force, Jun. 2004.

[2] Y. Nozaki, H. Tuncer, and N. Shenoy, "A tiered addressing scheme based on a floating cloud internetworking model," in *Proceedings of the 12th international conference on Distributed computing and networking*, ser. ICDCN'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 382–393.

[3] N. Shenoy, M. Yuksel, A. Gupta, K. Kar, V. Perotti, and M. Karir, "RAIDER: Responsive architecture for inter-domain economics and routing," in *GLOBECOM Workshops (GC Wkshps), 2010 IEEE*, Dec 2010, pp. 321–326.

[4] National Science Foundation Future Internet Design Initiative. [Online]. Available: http://www.nets-fia.net/

[5] National Science Foundation Future Internet Architecture Project. [Online]. Available: http://www.nets-fia.net/

[6] European Future Internet Portal. [Online]. Available: http://www.future-internet.eu/

[7] Asia Future Internet Forum. [Online]. Available: http://www.asiafi.net/

[8] New Generation Network Research Center. [Online]. Available: http://www2.nict.go.jp/w/w100/index-e

[9] Global Environment for Network Innovations (GENI). [Online]. Available: http://www.geni.net/

[10] Future Internet Research and Experimentation. [Online]. Available: http://cordis.europa.eu/fp7/ict/fire/

[11] I. Akyildiz, J. Xie, and S. Mohanty, "A survey of mobility management in next-generation all-IP-based wireless systems," *Wireless Communications, IEEE*, vol. 11, no. 4, pp. 16–28, Aug. 2004.

[12] D. Saha, A. Mukherjee, I. Misra, and M. Chakraborty, "Mobility support in IP: a survey of related protocols," *Network, IEEE*, vol. 18, no. 6, pp. 34–40, Nov 2004.

[13] P. Reinbold and O. Bonaventure, "IP micro-mobility protocols," *Communications Surveys Tutorials, IEEE*, vol. 5, no. 1, pp. 40–57, quarter 2003.

[14] D. H. Jun-Zhao Sun and J. Sauvola, "Mobility management techniques for the next generation wireless networks," ser. APOC 2001: Asia-Pacific optical and wireless communications. Bellingham, WA, USA: Society of Photo-Optical Instrumentation Engineers, 2001, pp. 155–166.

[15] I. AI-Surmi, M. Othman, and B. M. Ali, "Review on mobility management for Future-IP-based next generation wireless networks," pp. 989–994, 2010.

[16] L. J. Zhang, L. Zhang, L. Marchand, and S. Pierre, "A Survey of IP Layer Mobility Management Protocols in Next Generation Wireless Networks," *Next Generation Mobile Networks and Ubiquitous Computing*, pp. 79–93, 2010.

[17] J. Xie and X. Wang, "A Survey of Mobility Management in Hybrid Wireless Mesh Networks," *Network, IEEE*, vol. 22, no. 6, pp. 34–40, Nov 2008.

[18] F. Abduljalil and S. Bodhe, "A survey of integrating IP mobility protocols and mobile ad hoc networks," *Communications Surveys Tutorials, IEEE*, vol. 9, no. 1, pp. 14–30, quarter 2007.

[19] Y. P. Yang Xiao, Kin K. Leung and X. Du, "Architecture, mobility management, and quality of service for integrated 3G and WLAN networks," *Wireless Communications and Mobile Computing*, vol. 5, no. 7, pp. 805–823, Nov 2005.

[20] G. Lampropoulos, N. Passas, L. Merakos, and A. Kaloxylos, "Handover management architectures in integrated WLAN/cellular networks," *Communications Surveys Tutorials, IEEE*, vol. 7, no. 4, pp. 30–44, quarter 2005.

[21] A. Rahaman, J. Abawajy, and M. Hobbs, "Taxonomy and Survey of Location Management Systems," in *Computer and Information Science, 2007. ICIS 2007. 6th IEEE/ACIS International Conference on*, Jul 2007, pp. 369–374.

[22] T. Miyata, Y. Koga, P. Madsen, S.-I. Adachi, Y. Tsuchiya, Y. Sakamoto, and K. Takahashi, "A Survey on Identity Management Protocols and Standards," *IEICE - Trans. Inf. Syst.*, vol. E89-D, pp. 112–123, Jan 2006.

[23] P. Zhang, A. Durresi, and L. Barolli, "A Survey of Internet Mobility," in *Network-Based Information Systems, 2009. NBIS '09. International Conference on*, Aug 2009, pp. 147–154.

[24] M. Conti, S. Chong, S. Fdida, W. Jia, H. Karl, Y.-D. Lin, P. Mhnen, M. Maier, R. Molva, S. Uhlig, and M. Zukerman, "Research challenges towards the Future Internet," *Computer Communications*, vol. 34, no. 18, pp. 2115–2134, 2011.

[25] J. Roberts, "The clean-slate approach to future Internet design: a survey of research initiatives," *Annals of Telecommunications*, vol. 64, pp. 271–276, 2009.

[26] (2011) Future Internet Assembly 2011: Achievements and Technological Promises.

[27] N. Ekiz, T. Salih, S. Kucukoner, and K. Fidanboylu, "An Overview of Handoff Techniques in Cellular Networks," *International Journal Of Information Technology*, vol. 2, 2005.

[28] A. De La Oliva, A. Banchs, I. Soto, T. Melia, and A. Vidal, "An overview of IEEE 802.21: media-independent handover services," *Wireless Communications, IEEE*, vol. 15, no. 4, pp. 96 –103, aug. 2008.

[29] C. Perkins, "IP Mobility Support for IPv4," RFC 3344 (Proposed Standard), Internet Engineering Task Force, Aug. 2002, obsoleted by RFC 5944, updated by RFC 4721.

[30] N. Moore, "Optimistic Duplicate Address Detection (DAD) for IPv6," RFC 4429 (Proposed Standard), Internet Engineering Task Force, Apr. 2006.

[31] C. Perkins, "IP Encapsulation within IP," RFC 2003 (Proposed Standard), Internet Engineering Task Force, Oct. 1996, updated by RFC 3168.

[32] H. Levkowetz and S. Vaarala, "Mobile IP Traversal of Network Address Translation (NAT) Devices," RFC 3519 (Proposed Standard), Internet Engineering Task Force, Apr. 2003.

[33] J. Arkko, C. Vogt, and W. Haddad, "Enhanced Route Optimization for Mobile IPv6," RFC 4866 (Proposed Standard), Internet Engineering Task Force, May 2007.

[34] R. Koodli, "Mobile IPv6 Fast Handovers," RFC 5568 (Proposed Standard), Internet Engineering Task Force, Jul. 2009.

[35] H. Soliman, C. Castelluccia, K. ElMalki, and L. Bellier, "Hierarchical Mobile IPv6 (HMIPv6) Mobility Management," RFC 5380 (Proposed Standard), Internet Engineering Task Force, Oct. 2008.

[36] S. Thomson, T. Narten, and T. Jinmei, "IPv6 Stateless Address Autoconfiguration," RFC 4862 (Draft Standard), Internet Engineering Task Force, Sep. 2007.

[37] Network-based Localized Mobility Management IETF Working Group. [Online]. Available: http://datatracker.ietf.org/wg/netlmm/charter/

[38] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, "Proxy Mobile IPv6," RFC 5213 (Proposed Standard), Internet Engineering Task Force, Aug. 2008.

[39] K. Kong, W. Lee, Y.-H. Han, M.-K. Shin, and H. You, "Mobility management for all-IP mobile networks: mobile IPv6 vs. proxy mobile IPv6," *Wireless Communications, IEEE*, vol. 15, no. 2, pp. 36–45, Apr 2008.

[40] J. Korhonen and V. Devarapalli, "Local Mobility Anchor (LMA) Discovery for Proxy Mobile IPv6," RFC 6097 (Informational), Internet Engineering Task Force, Feb. 2011.

[41] I. Al-Surmi, M. Othman, and B. Ali, "Review on mobility management for future-IP-based next generation wireless networks," in *Advanced Communication Technology (ICACT), 2010 The 12th International Conference on*, vol. 2, feb. 2010, pp. 989 –994.

[42] J.-H. Lee, T. Ernst, and T.-M. Chung, "Cost analysis of IP mobility management protocols for consumer mobile devices," *Consumer Electronics, IEEE Transactions on*, vol. 56, no. 2, pp. 1010 –1017, may 2010.

[43] J.-I. Kim, H. Jung, and S. J. Koh, "Distributed mobility control for mobile-oriented Future Internet environments," in *ICT Convergence (ICTC), 2011 International Conference on*, sept. 2011, pp. 342 –347.

[44] MobilityFirst Future Internet Architecture Project. [Online]. Available: http://mobilityfirst.winlab.rutgers.edu/

[45] S. C. Nelson, G. Bhanage, and D. Raychaudhuri, "GSTAR: generalized storage-aware routing for mobilityfirst in the future mobile internet," in *Proceedings of the sixth international workshop on MobiArch*, ser. MobiArch '11. New York, NY, USA: ACM, 2011, pp. 19–24.

[46] S. Paul, R. Yates, D. Raychaudhuri, and J. Kurose, "The cache-and-forward network architecture for efficient mobile content delivery services in the future internet," in *Innovations in NGN: Future Network and Services, 2008. K-INGN 2008. First ITU-T Kaleidoscope Academic Conference*, May 2008, pp. 367–374.

[47] L. Subramanian, M. Caesar, C. T. Ee, M. Handley, M. Mao, S. Shenker, and I. Stoica, "HLP: a next generation inter-domain routing protocol," *SIGCOMM Comput. Commun. Rev.*, vol. 35, pp. 13–24, Aug 2005.

[48] Expressive Internet Architecture. [Online]. Available: http://www.cs.cmu.edu/ xia/

[49] A. Anand, F. Dogar, D. Han, B. Li, H. Lim, M. Machadoy, W. Wu, A. Akella, D. Andersen, J. Byersy, S. Seshan, and P. Steenkiste, "XIA: An Architecture for an Evolvable and Trustworthy Internet," Department of Computer Science, Carnegie Mellon- University, Tech. Rep. Technical Report CMU-CS-11-100, Feb 2011.

[50] F. R. Dogar and P. Steenkiste, "Segment based Inter-networking to Accommodate Diversity at the Edge," Department of Computer Science, Carnegie Mellon- University, Tech. Rep. CMU CSD technical report, CMU-CS-10-104, Feb 2010.

[51] N. Niebert, M. Prytz, A. Schieder, N. Papadoglou, L. Eggert, F. Pittmann, and C. Prehofer, "Ambient networks: a framework for future wireless internetworking," in *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, vol. 5, May 2005, pp. 2969–2973.

[52] R. Calvo, A. Surtees, J. Eisl, and M. Georgiades, "Mobility Management in Ambient Networks," in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, Apr 2007, pp. 894–898.

[53] J. Laganier and A. R. Prasad, "Interactions of ambient networks composition and mobility toolbox frameworks," in *Proceedings of the 3rd international conference on Mobile technology, applications & systems*, ser. Mobility '06. New York, NY, USA: ACM, 2006.

[54] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol," RFC 3261 (Proposed Standard), Internet Engineering Task Force, Jun. 2002, updated by RFCs 3265, 3853, 4320, 4916, 5393, 5621, 5626, 5630, 5922, 5954, 6026, 6141.

[55] R. Stewart, "Stream Control Transmission Protocol," RFC 4960 (Proposed Standard), Internet Engineering Task Force, Sep. 2007, updated by RFC 6096.

[56] DAIDALOS, Designing Advanced network Interfaces for the Delivery and Administration of Location independent, Optimised personal Services project. [Online]. Available: http://www.ist-daidalos.org/

[57] E. Papadopoulou, S. McBurney, N. Taylor, and M. H. Williams, "Linking Privacy and User Preferences in the Identity Management for a Pervasive System," in *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01*. Washington, DC, USA: IEEE Computer Society, 2008, pp. 192–195.

[58] S. Sargento and R. Sarro, "Architecture and Design: Routing for Ad-hoc and Moving Networks," Tech. Rep. DII-241, Oct. 2007.

[59] R. Moskowitz, P. Nikander, P. Jokela, and T. Henderson, "Host Identity Protocol," RFC 5201 (Experimental), Internet Engineering Task Force, Apr. 2008.

[60] R. Moskowitz and P. Nikander, "Host Identity Protocol (HIP) Architecture," RFC 4423 (Informational), Internet Engineering Task Force, May 2006.

[61] A. Gurtov, *Host Identity Protocol (HIP): Towards the Secure Mobile Internet*. WILEY, Aug 2008, no. ISBN: 978-0-470-99790-1.

[62] Host Identity Protocol (HIP) IETF Work Group. [Online]. Available: http://www.ietf.org/dyn/wg/charter/hip-charter.html

[63] T. Henderson, "End-Host Mobility and Multihoming with the Host Identity Protocol," Internet draft, Internet Engineering Task Force, Jun. 2006.

[64] P. Nikander, J. Ylitalo, and J. Wall, "Integrating security, mobility, and multi-homing in a hip way," in *NDSS 2003 Proceedings of the Network and Distributed Systems Security Symposium*, Feb. 2003, pp. 87–99.

[65] P. Nikander, T. Henderson, C. Vogt, and J. Arkko, "End-Host Mobility and Multihoming with the Host Identity Protocol," RFC 5206 (Experimental), Internet Engineering Task Force, Apr. 2008.

[66] J. Laganier, "Host Identity Protocol (HIP) Randezvous Extention," Internet draft, Internet Engineering Task Force, Oct. 2005.

[67] S. Novaczki, L. Bokor, and S. Imre, "Micromobility support in HIP: survey and extension of host identity protocol," in *Electrotechnical Conference, 2006. MELECON 2006. IEEE Mediterranean*, may 2006, pp. 651–654.

[68] M. Muslam, H. Chan, and N. Ventura, "Host Identity Protocol extension supporting localized mobility management," in *Consumer Communications and Networking Conference (CCNC), 2011 IEEE*, Jan 2011, pp. 106–110.

[69] N. I. of Information and C. T. N. of Japan, "New Generation Network Architecture AKARI Conceptual Design v 1.1," Tech. Rep., Oct 2008. [Online]. Available: http://akari-project.nict.go.jp/eng/index2.htm

[70] i3 project. [Online]. Available: http://i3.cs.berkeley.edu/

[71] I. Stoica, D. Adkins, S. Zhuang, S. Shenker, and S. Surana, "Internet Indirection Infrastructure," *Networking, IEEE/ACM Transactions on*, vol. 12, no. 2, pp. 205–218, apr 2004.

[72] S. Zhuang, K. Lai, I. Stoica, R. Katz, and S. Shenker, "Host mobility using an internet indirection infrastructure," *Wirel. Netw.*, vol. 11, pp. 741–756, Nov 2005.

[73] A. Gurtov, D. Korzun, A. Lukyanenko, and P. Nikander, "Hi3: An efficient and secure networking architecture for mobile hosts," *Comput. Commun.*, vol. 31, pp. 2457–2467, Jun 2008.

[74] D. Korzun and A. Gurtov, "On scalability properties of the Hi3 control plane," *Comput. Commun.*, vol. 29, pp. 3591–3601, Nov 2006.

[75] P. Nikander and J. Arkko, "Host Identity Indirection Infrastructure (Hi3)," Internet draft, Internet Engineering Task Force Network WG, Jun. 2004.

[76] Locator/id separation protocol (lisp) ietf work group. [Online]. Available: http://www.ietf.org/dyn/wg/charter/lisp-charter.html

[77] D. Farinacci, "Locator/ID Separation Protocol (LISP)," Internet draft, Internet Engineering Task Force Network WG, Sep. 2009.

[78] R. Hinden, "New Scheme for Internet Routing and Addressing (ENCAPS) for IPNG," RFC 1955 (Informational), Internet Engineering Task Force, Jun. 1996.

[79] D. Lewis, "Interworking LISP with IPv4 and IPv6," Internet draft, Internet Engineering Task Force, May 2009.

[80] V. Fuller, "LISP MAP Server," Internet draft, Internet Engineering Task Force, Oct. 2009.

[81] V. Fuller, "LISP Alternative Topology (LISP+ALT)," Internet draft, Internet Engineering Task Force, May 2009.

[82] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," Proposed Standard, Internet Engineering Task Force, 1995.

[83] D. Meyer, "The Locator Identifier Separation Protocol (LISP)," *Cisco-The Internet Protocol*, vol. 11.

[84] J. Pan, S. Paul, R. Jain, and M. Bowman, "MILSA: A Mobility and Multihoming Supporting Identifier Locator Split Architecture for Naming in the Next Generation Internet," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, Dec 2008, pp. 1–6.

[85] J. Pan, S. Paul, R. Jain, and X. Xu, "Hybrid Transition Mechanism for MILSA Architecture for the Next Generation Internet," in *GLOBECOM Workshops, 2009 IEEE*, Dec 2009, pp. 1–6.

[86] R. Jian, "Internet 3.0: Ten Problems with Current Internet Architecture and Solutions for the Next Generation," in *Military Communications Conference, 2006. MILCOM 2006. IEEE*, Oct 2006, pp. 1–9.

[87] J. Pan, R. Jain, S. Paul, M. Bowman, X. Xu, and S. Chen, "Enhanced MILSA Architecture for Naming, Addressing, Routing and Security Issues in the Next Generation Internet," in *Communications, 2009. ICC '09. IEEE International Conference on*, Jun 2009, pp. 1–6.

[88] CARMEN, CARrier grade MEsh Networks. [Online]. Available: http://www.ict-carmen.eu/

[89] P. Patras, "CARMEN,ratified architecture deliverable," Tech. Rep. D1.2., Jan. 2009.

[90] "IEEE Draft Standard for Local and Metropolitan Area Networks: Media Independent Handover Services," *IEEE Unapproved Draft Std P802.21/D14, Sept. 2008*, 2008.

[91] H. Marques, J. Rodriguez, and P. Marques, "D3.2: Design of optimized handover operations for heterogeneous wireless systems," Tech. Rep., Jan. 2008. [Online]. Available: http://www.ict-hurricane.eu/

[92] "IEEE Standard for Local and Metropolitan Area Networks- Part 21: Media Independent Handover," *IEEE Std 802.21-2008*, pp. 1–301, 2009.

[93] "IEEE Standard for Architectural Building Blocks Enabling Network-Device Distributed Decision Making for Optimized Radio Resource Usage in Heterogeneous Wireless Access Networks," *IEEE Std 1900.4-2009*, pp. 1–119, 2009.

[94] M. Buddhikot, A. Hari, K. Singh, and S. Miller, "MobileNAT: a new technique for mobility across heterogeneous address spaces," *Mobile Networks and Applications*, vol. 10, pp. 289–302, Jun 2005.

[95] J. Kempf, "Problem Statement for Network-Based Localized Mobility Management (NETLMM)," RFC 4830 (Informational), Internet Engineering Task Force, Apr. 2007.

[96] L. Osborne, A. Abdel-Hamid, and R. Ramadugu, "A performance comparison of Mobile IPv6, hierarchical Mobile IPv6, and Mobile IPv6 regional registrations," in *Wireless Networks, Communications and Mobile Computing, 2005 International Conference on*, vol. 2, Jun 2005, pp. 1545–1550.

[97] X. Pérez-Costa, M. Torrent-Moreno, and H. Hartenstein, "A performance comparison of Mobile IPv6, Hierarchical Mobile IPv6, fast handovers for Mobile IPv6 and their combination," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 7, pp. 5–19, Oct 2003.

[98] N. Jordan, A. Poropatich, and J. Fabini, "Performance evaluation of the hierarchical Mobile IPv6 approach in a WLAN hotspot scenario," in *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, vol. 5, May 2005, pp. 2810–2814.

[99] M.-K. Yi, J.-W. Choi, and Y.-K. Yang, "A Comparative Analysis on the Signaling Load of Proxy Mobile IPv6 and Hierarchical Mobile IPv6," in *Wireless Pervasive Computing, 2009. ISWPC 2009. 4th International Symposium on*, Feb 2009, pp. 1–5.

[100] J.-H. Lee, T.-M. Chung, and S. Gundavelli, "A comparative signaling cost analysis of Hierarchical Mobile IPv6 and Proxy Mobile IPv6," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, Sep 2008, pp. 1–6.

[101] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research," *WIRELESS COMMUNICATIONS & MOBILE COMPUTING (WCMC): SPECIAL ISSUE ON MOBILE AD HOC NETWORKING: RESEARCH, TRENDS AND APPLICATIONS*, vol. 2, pp. 483–502, 2002.

[102] F. Bai and A. Helmy, "Chapter 1: A Survey of Mobility Models in Wireless Adhoc Networks:," *Wireless Ad hoc and Sensor Networks*, vol. 2, no. 5, pp. 1–30, 2006.

[103] M. Musolesi and C. Mascolo, *Mobility Models for Systems Evaluation*, 2009, p. 43.

[104] I. Akyildiz and W. Wang, "A dynamic location management scheme for next-generation multitier pcs systems," *Wireless Communications, IEEE Transactions on*, vol. 1, no. 1, pp. 178 –189, jan 2002.

[105] J. Ariyakhajorn, P. Wannawilai, and C. Sathitwiriyawong, "A comparative study of random waypoint and gauss-markov mobility models in the performance evaluation of manet," in *Communications and Information Technologies, 2006. ISCIT '06. International Symposium on*, 2006, pp. 894–899.

[106] F. Bai, N. Sadagopan, and A. Helmy, "IMPORTANT: a framework to systematically analyze the Impact of Mobility on Performance of Routing Protocols for Adhoc Networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 2, march-3 april 2003, pp. 825 – 835 vol.2.

[107] J. Harri, F. Filali, and C. Bonnet, "Mobility models for vehicular ad hoc networks: a survey and taxonomy," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 4, pp. 19–41, 2009.

[108] R. Roy, "Fluid-flow mobility," in *Handbook of Mobile Ad Hoc Networks for Mobility Models*. Springer US, 2011, pp. 405–441.

[109] S. Pack, Y. Choi, and M. Nam, "Design and analysis of optimal multi-level hierarchical mobile ipv6 networks," *Wirel. Pers. Commun.*, vol. 36, pp. 95–112, January 2006.

[110] J.-H. Lee, Y.-H. Han, S. Gundavelli, and T.-M. Chung, "A comparative performance analysis on hierarchical mobile ipv6 and proxy mobile ipv6," *Telecommunication Systems*, vol. 41, pp. 279–292, 2009.

[111] M. Afanasyev, T. Chen, G. Voelker, and A. Snoeren, "Usage patterns in an urban wifi network," *Networking, IEEE/ACM Transactions on*, vol. 18, no. 5, pp. 1359–1372, 2010.

[112] CRAWDAD,A Community Resource for Archiving Wireless Data At Dartmouth. [Online]. Available: http://www.crawdad.org/

[113] Y.-B. Lin, "Reducing location update cost in a pcs network," *Networking, IEEE/ACM Transactions on*, vol. 5, no. 1, pp. 25–33, 1997.

[114] J. Xie and I. Akyildiz, "A distributed dynamic regional location management scheme for mobile ip," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, 2002, pp. 1069–1078 vol.2.

[115] A. Vilhar, R. Novak, and G. Kandus, "The impact of network topology on the performance of MAP selection algorithms," *Comput. Netw.*, vol. 54, pp. 1197–1209, May 2010.

[116] S. Jeon and Y. Kim, "Adaptive handoff management in the proxy mobile ipv6 domain," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011, pp. 1–5.

[117] Y.-H. Han, J. Choi, and S.-H. Hwang, "Reactive Handover Optimization in IPv6-Based Mobile Networks," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 9, pp. 1758 –1772, sept. 2006.

[118] A. Dutta, T. Zhang, Y. Ohba, K. Taniuchi, and H. Schulzrinne, "MPA assisted Optimized Proactive Handoff Scheme," in *Proceedings of The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 155–165.

[119] J. Schumacher, M. Dobler, E. Dillon, G. Power, M. Fiedler, D. Erman, K. De Vogeleer, M. Ramos, and J. Argente, "Providing an User Centric Always Best Connection," in *Evolving Internet (INTERNET), 2010 Second International Conference on*, September 2010, pp. 80 –85.

[120] A. Calvagna and G. Di Modica, "A user-centric analysis of vertical handovers," in *Proceedings of the 2nd ACM international workshop on Wireless mobile applications and services on WLAN hotspots*, ser. WMASH '04. New York, NY, USA: ACM, 2004, pp. 137–146.

[121] S. Islam and J.-C. Grégoire, "User-centric service provisioning for IMS," in *Proceedings of the 6th International Conference on Mobile Technology, Application & Systems*, ser. Mobility '09. New York, NY, USA: ACM, 2009, pp. 5:1–5:8.

[122] A. Eriksson and B. Ohlman, "Dynamic Internetworking Based on Late Locator Construction," in *IEEE Global Internet Symposium, 2007*, may 2007, pp. 67 –72.

[123] X. Yang, D. Clark, and A. Berger, "NIRA: A New Inter-Domain Routing Architecture," *Networking, IEEE/ACM Transactions on*, vol. 15, no. 4, pp. 775 –788, aug. 2007.

[124] Y. Nozaki, H. Tuncer, and N. Shenoy, "Isp tiered model based architecture for routing scalability," in *Communications (ICC), 2012 IEEE International Conference on*, 2012, pp. 5817–5821.

[125] Yong Li and Haibo Su and Li Su and Depeng Jin and Lieguang Zeng, "A Comprehensive Performance Evaluation of PMIPv6 over IP-Based Cellular Networks," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, april 2009, pp. 1 –6.

[126] C. Makaya and S. Pierre, "An Analytical Framework for Performance Evaluation of IPv6-Based mobility Management Protocols," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 3, pp. 972 –983, march 2008.

[127] T. Narten, E. Nordmark, W. Simpson, and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)," RFC 4861 (Draft Standard), Internet Engineering Task Force, Sep. 2007, updated by RFC 5942.

[128] Y.-H. Han and S.-H. Hwang, "Movement detection analysis in mobile IPv6," *Communications Letters, IEEE*, vol. 10, no. 1, pp. 59 – 61, jan 2006.

[129] Z. Liang, J. Zheng, S. Ma, S. Yang, B. Wang, and J. Liu, "Modeling and analysis of queuing effect on Mobile IPv6 handoff performance," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, sept. 2010, pp. 2282 –2287.

[130] J. Xie, I. Howitt, and I. Shibeika, "IEEE 802.11-Based Mobile IP Fast Handoff Latency Analysis," in *Communications, 2007. ICC '07. IEEE International Conference on*, june 2007, pp. 6055 –6060.

[131] H. Fathi, S. S. Chakraborty, and R. Prasad, "Optimization of Mobile IPv6-Based Handovers to Support VoIP Services in Wireless Heterogeneous Networks," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 1, pp. 260 –270, jan. 2007.

[132] R. Droms, J. Bound, B. Vloz, T. Lemon, C. Perkins, and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)," RFC 3315 (Standards Track), Internet Engineering Task Force, Jul. 2003.

[133] J. Laganier, S. Narayanan, and P. McCann, "Interface between a Proxy MIPv6 Mobility Access Gateway and a Mobile Node," Internet draft, Internet Engineering Task Force, Feb. 2008.

[134] F. F. Liza and W. Yao, "Implementation Architecture of Proxy Mobile IPv6 Protocol for NS2 Simulator Software," in *Proceedings of the 2009 International Conference on Communication Software and Networks*, ser. ICCSN '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 287–291.

[135] H. Jeong, S. Maeng, and Y. Chae, "HIMIPv6: An Efficient IP Mobility Management Protocol for Broadband Wireless Networks," *IEICE Transactions on Information and Systems*, vol. E92.D, no. 10, pp. 1857–1866, 2009.

[136] Z. Mao and C. Douligeris, "A distributed database architecture for global roaming in next-generation mobile networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 1, pp. 146–160, Feb. 2004.

[137] H. Soliman, *Mobile IPv6: mobility in a wireless Internet*. Addison-Wesley, April 2004.

[138] J.-H. Lee, T. Ernst, and T.-M. Chung, "Cost analysis of ip mobility management protocols for consumer mobile devices," *IEEE Trans. on Consum. Electron.*, vol. 56, no. 2, pp. 1010–1017, May 2010. [Online]. Available: http://dx.doi.org/10.1109/TCE.2010.5506033

[139] J. Gross and J. Yellen, *Handbook of Graph Theory*, ser. Discrete Mathematics and Its Applications. Taylor & Francis, 2004. [Online]. Available: http://books.google.com/books?id=mKkIGIea_BkC

[140] H. Tuncer, A. Kwasinski, and N. Shenoy, "Performance analysis of virtual mobility domain scheme vs. {IPv6} mobility protocols," *Computer Networks*, vol. 57, no. 13, pp. 2578 – 2596, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128613001448

[141] H. Tuncer, Y. Nozaki, and N. Shenoy, "Virtual Domains for Seamless User Mobility," in *Proceedings of the 9th ACM international symposium on Mobility management and wireless access*, ser. MobiWac'11. ACM, 2011.

[142] ——, "Virtual mobility domains - a mobility architecture for the future internet," in *Communications (ICC), 2012 IEEE International Conference on*, 2012, pp. 2774–2779.

[143] S. Seshan, H. Balakrishnan, and R. H. Katz, "Handoffs in cellular wireless networks: The daedalus implementation and experience," *Wireless Personal Networks*, vol. 4, no. 2, pp. 141–162, Jan. 1997.

[144] Description of solvers available in GAMS. [Online]. Available: http://www.gams.com/solvers/solvers.htm

[145] H. Anderson, ""metropolis, monte carlo and the maniac",," in *"Los Alamos Science 14"*, 1986, p. 96108.

[146] M. R. Bussieck and A. Pruessner, "Mixed-Integer Nonlinear Programming," *Communications Surveys Tutorials, IEEE*, vol. SIA/OPT Newsletter: Views and News, February 2003.

[147] DICOPT (DIscrete and Continuous OPTimizer). [Online]. Available: http://www.gams.com/dd/docs/solvers/dicopt.pdf

[148] 802.11ac: The fifth generation of wi-fi. [Online]. Available: http://www.cisco.com/c/en/us/products/collateral/wireless/aironet-3600-series/white_paper_c11-713103.pdf

[149] The Mobile Broadband Standard - LTE Advanced. [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced